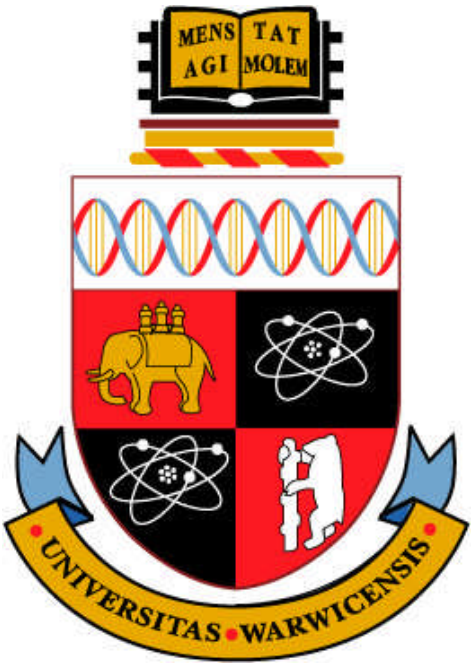


Mathematical models characterising the kinetics and dynamics of topotecan to account for drug resistance mechanisms

by

Mohammed Isam Atari

Thesis submitted in partial fulfilment of the requirements for admission to the degree of

Doctor of Philosophy in Engineering

School of Engineering

January 2011

THE UNIVERSITY OF
WARWICK

I dedicate this thesis to the memory of my beloved father, Isam Atari and my grandmother Fathiyyeh Abdulrahim

Table of Contents

List of Figures

List of Tables

Abbreviations

ABC	Adenosine triphosphate-binding cassette
ADME	Absorption, Distribution, Metabolism and Elimination
APC	Anaphase-promoting complex
a.u.	Arbitrary units
ATP	Adenosine triphosphate
ALDH	Aldehyde dehydrogenase
ATM	Ataxia telangiectasia mutated
AUC	Area under the curve
BCRP	Breast cancer resistance protein
CAK	CDK-activating kinase
CCPM	Cell cycle phase marker
CDC2	Cell division cycle
cDNA	Complementary DNA
CPT	Camptothecin
CDK	Cyclin-dependent kinase
CKI	Cyclin-dependent kinase inhibitor
CORI	Correlation index for the residuals
$CORI_{mean}$	Mean absolute correlation index
CRS	Cytoplasmic retention signal
DNA	Deoxyribonucleic acid
dsDNA	double-stranded DNA
G_1	Gap 1
G_2	Gap 2

γH2AX	Gamma-H2AX
GFP	Green fluorescent protein
H	Hours
HPLC	High performance liquid chromatography
HTC	High-throughput screen
IMT	Inter-mitotic time
M	Mitotic
MSD	Membrane spanning domain
MXR	Mitoxantrone-resistance protein
MRP1	Multidrug resistance protein 1
Min	Minutes
M	Molar
NBD	Nucleotide binding domain
NWD	Not well-determined
ODE	Ordinary differential equation
P-gp	P-glycoprotein
PP2C	Type 2C protein phosphatise
ABCP	Placenta-specific ABC protein
PSSA	Pseudo-steady state approximation
PBS	Phosphate buffered saline
PD	Pharmacodynamics
PK	Pharmacokinetics
PBPK/PD	physiologically-based PK/PD
ROI	Regions-of-interest
RNA	Ribonucleic acid
RSS	Residual sum of squares

S	Synthesis
SDLN	standard deviation of the natural logarithm
ssDNA	single-stranded DNA
TPLSM	Two-photon laser-scanning microscopy
TPT	Topotecan
TPT_H	Topotecan hydroxy acid form
TPT_L	Topotecan lactone form
T161	Threonine 161
Y15	Tyrosine 15
S	Seconds

Acknowledgements

Certainly the greatest aspect of being a postgraduate student and working towards a Ph.D. is the people you worked with. First and foremost, I would like to thank my project supervisors, Dr. Neil D. Evans and Dr. Michael J. Chappell for their guidance, their invaluable criticism, and their willingness to spend much time supervising my project. They have both been extremely helpful in simplifying and clarifying pitfalls in mathematical analyses. In addition, I am indebted to them for the important and inspiring ideas that they brought up throughout this work that have helped me focus my attention on wider areas relating to this project. Generally, this process has allowed me to acquire more experience than I could ever imagine.

I am indeed grateful to the University of Warwick for offering me the Vice Chancellor's Ph.D. Scholarship which provided me with the financial support.

Much of the work in this thesis was done in the context of the long term study between the University of Warwick and Wales College of Medicine (Cardiff University), on the effect that anti-cancer agents have on a certain type of cells. I would like to show my gratitude to the whole team, especially Professor Paul J. Smith, Dr. Rachel J. Errington, and Dr. Imtiaz Khan for providing me with two-photon scanning-laser microscopy data from live human breast cancer cells (MCF-7 cell line), high-resolution fluorescence data from live human osteosarcoma cells (U-2 OS cell line) and DNA damage data (γH2AX) in addition to them explaining to me the microscopy techniques carried out at Cardiff University. They helped to make biology fun for me.

I am also thankful for the many fun times spent with my office mates in A405B: Szevone Chin, Thomas Grandjean and Mary Finnegan. I would like to express my thanks to the support staff in the School of Engineering at the University of Warwick and to Ms. Kerrie Hatton in particular for assisting me in many different ways.

I owe my deepest gratitude to my mother Samar, my grandfather Sudqi Abdulrahim and my uncle Samir Abdulrahim. They raised me, supported me, taught me and loved me. I

would like to thank my siblings, Ali, Alia, Lina and Linda for their immense love and for giving me the opportunity one could ask for. Their frequent support and encouragement is a blessing in my life.

Lastly and most importantly, I wish to thank Beesan, the love of my life, without her understanding, encouragement and diligence, none of this would be possible.

Declarations

This thesis is the original work of the author, with the following publications describing parts of the work:

Chapter 2:

M. I. Atari, R. Ali, M. J. Chappell and N. D. Evans, "The Role of Systems Modelling in Biomedical Engineering", In: Proceedings of the 1st International Symposium on Medical Informatics and Biomedical Engineering, Amman, Jordan, March 20-22, 2006.

Chapter 4:

Mohammed I. Atari, Michael J. Chappell, Rachel J. Errington, Paul J. Smith and Neil D. Evans, "Kinetic modelling of the role of the aldehyde dehydrogenase enzyme and the breast cancer resistance protein in drug resistance and transport", Computer Methods and Programs in Biomedicine. DOI: 10.1016/j.cmpb.2010.06.008

Mohammed I. Atari, Michael J. Chappell, Rachel J. Errington, Paul J. Smith and Neil D. Evans, "Kinetic modelling of the role of the aldehyde dehydrogenase enzyme and the breast cancer resistance protein in drug resistance and transport", Submitted to: The 7th IFAC Symposium on Modelling and Control in Biomedical Systems (including Biological Systems), Aalborg, Denmark, August 12-14, 2009.

Summary

This thesis describes the use of mathematical modelling in studying the kinetics and dynamics of the anti-cancer agent topotecan (TPT), a semi-synthetic derivative of the natural extract camptothecin (CPT), which has been found to act as an inhibitor of the DNA enzyme topoisomerase I in a specific and reversible fashion. The drug undergoes reversible hydrolysis from the pharmacologically active parent lactone form (TPT_L) to an inactive hydroxy acid form (TPT_H). In the cytoplasm the irreversible inactivation of TPT_L is catalysed by the enzyme aldehyde dehydrogenase (ALDH). Over-expression of the human breast cancer resistance protein (BCRP/ABCG2) has been linked to high levels of resistance to the anti-cancer agent TPT by promoting an active efflux pump mechanism. The expressions of both ALDH and BCRP have been experimentally identified in a large number of solid tumours and thus play an important role in clinical drug resistance of cancers. To investigate the catalytic reaction and efflux pump mechanism, a state-space model for the *in vitro* uptake kinetics of TPT has been extended to better describe the drug activity and delivery of TPT_L to the DNA target as well as the catalysis by ALDH and the elimination of drug from the cytoplasm via the efflux pump. All unknown model parameters were uniquely estimated to a high level of confidence. Model simulations have been compared with live human breast cancer cells data and found to give good qualitative agreement. In addition, a cell cycle model has been extended to include the inhibition effect of the protein $p21^{CIP1/WAF1}$ on the cell cycle traverse and the kinetic model has then been linked to the cell cycle model, which facilitates analysis of the response of the growth of single cells in the presence and absence of TPT. All unknown model parameters were uniquely determined by the output structure corresponding to the experiment. Parameter estimation was performed using green fluorescent protein tagged Cyclin B1 data for the osteosarcoma cell line U-2 OS. The *novel* coupling of both models allows the study of drug perturbation to the cell cycle as well as *in silico* estimation and prediction of the relationship between the target binding and the dose, also permitting the effects of different levels of expression of the drug resistance protein and the ALDH enzyme. Such a coupled kinetic/dynamic model, once fully validated, has the potential for enhancing the design of optimal dosing regimens.

1 Introduction

1.1 Problem statement and motivation

Drug discovery and development is a time-consuming and expensive process, with very few novel drug compounds getting approved [1, 2]. The cost of bringing a new drug to the market is about $1 billion and the overall development process takes an average of 12-15 years [3]. This process has a low success rate, particularly during clinical trials, due to reasons including lack of understanding of the pharmacokinetic[1]/pharmacodynamic[2] (dose-response) relationship in addition to unexpected safety events [4-6].

One of the major goals in pharmacology is the quantitative prediction of the response to drugs (i.e., exploring drug effects). In recent years, predicting and optimising drug kinetic/dynamic properties along the discovery process (i.e., *in vitro*) rather than at the final stages (i.e., *in vitro*), has become increasingly important. In addition, predictive modelling tools can provide key information to enhance and speed-up the drug discovery and development process. Therefore, several *in vitro* and *in silico* approaches have been devised to predict invaluable drug kinetic/dynamic properties [1] and to greatly assist the decision-making process [7].

In vitro drug kinetics/dynamics modelling establishes a mathematical and theoretical link between these two processes and helps better understand and predict drug action [8]. This modelling approach is being incorporated into several drug discovery and development projects and is having a growing impact [9-11]. Numerous anti-cancer agents require interaction with DNA or chromatin components of cancer cells to achieve therapeutic activity [12]. Accordingly, investigation and quantification of drug targeting dynamics can be greatly informative in the design of therapeutic regimens and in the field of drug discovery.

[1] Pharmacokinetics (PK): the study that characterises the absorption, distribution, metabolism and elimination (ADME) properties of a drug.
[2] Pharmacodynamics (PD): biological and physiological response to an administered drug.

The ultimate aim of this thesis is to develop a coupled (joint) mathematical model that is capable of describing the *in vitro* drug kinetics of the anti-cancer agent topotecan (TPT) for single live human osteosarcoma cells linked to primary biological responses (i.e., cell cycle). This coupling has been achieved by extending a previously published model by Evans *et al.* [13] for the *in vitro* uptake kinetics of TPT in order to account for cell transporters and drug resistance proteins. In addition, a cell cycle model has been extended from the model developed by Pomerening *et al.* [14] to include the inhibition effect of the protein $p21^{CIP1/WAF1}$ on the cell cycle process. The extended models (i.e., the drug kinetics model and the cell cycle model) have been linked to demonstrate both the dynamics and temporal interactions of the pharmacologically active form of the drug to its DNA-associated molecular target. By linking the models in this way, it is possible to demonstrate the response of the growth of single human cells in the presence and absence of the anti-cancer agent TPT.

1.2 Thesis structure

This thesis is divided into seven chapters. Chapter 2 contains the preliminary material on model development using state-space modelling techniques in addition to the analytical and numerical methods that will be used in this thesis. The mathematical tools discussed in Chapter 2 are fundamental in understanding the models, results and concluding remarks in Chapters 4, 5 and 6.

In Chapter 3 the background biology of the anti-cancer agent TPT, drug resistance mechanisms and a review of the cell cycle process shall be described to assist in building the mathematical models (in Chapters 4, 5 and 6) presented in this thesis. The biological background is not in great detail, however, it should be adequate for the reader to understand how the models were developed in the subsequent chapters.

The mathematical modelling begins in Chapter 4. The extended drug kinetics model is presented in this chapter (i.e., Chapter 4). The modelling starts with the basic building block for the full cell-based model presented in this thesis. That is, the reversible

hydrolysis of TPT which has been modelled by Evans *et al.* [15]. The following section in Chapter 4 describes the extended mathematical model for the *in vitro* uptake kinetics of TPT based on the assumptions related to drug resistance mechanisms discussed in Chapter 3. This is followed by structural identifiability analysis and parameter estimation techniques which are explained in Chapter 2. Parameter estimation in Chapter 4 was performed using data collected via two-photon laser-scanning microscopy (TPLSM) from live human breast cancer cells. These data were provided by the biological collaborators at Cardiff University. Heterogeneity across two different kinds of cells (a high loading cell and a low loading cell) has been addressed and compared to the average response of cells.

The extended cell cycle model for human cells is presented in Chapter 5. The cell cycle model described in Chapter 5 is a novel extension of a previously published cell model by Pomerening *et al.* [14] (for *Xenopus laevis* embryos) to describe the response of the growth of single human cells in the absence of the anti-cancer agent TPT. The novel extension of this model includes the inhibitory action of a protein involved in the cell cycle, the $p21^{CIP1/WAF1}$. This component has been added to the extended cell cycle model, based on the background biology discussed in Chapter 3. Similar to the mathematical techniques applied in Chapter 4, structural identifiability analysis of the extended cell cycle model was performed in addition to parameter estimation. Parameter estimation was performed using Cyclin B1 (cell cycle regulator) fluorescence data for single human osteosarcoma cells provided by the biological collaborators at Cardiff. In Chapter 5, the model was compared to live-cell data in the absence of TPT (i.e., for an untreated cell).

In Chapter 6, both the extended drug kinetics model (presented in Chapter 4) and the extended cell cycle model (presented in Chapter 4) are coupled by introducing an effect component (dynamics). This coupling is based on biological assumptions (given in Chapter 3) of the cell cycle perturbations induced by TPT. Before estimating the unknown model parameters a formal structural identifiability analysis was performed for the full drug kinetics/cell cycle response model. Similar to the type of data used in Chapter 5, parameter estimation (in the presence of two treatment regimens of TPT) was performed using live-cell Cyclin B1 fluorescence data in addition to DNA damage data. The results obtained from this Chapter were compared to the results from Chapter 5 (i.e., in the absence of TPT). To investigate how cells respond to the action of TPT.

To assist in comprehension, each chapter will close with a summary of the key results and findings in addition to the important features developed in it. In Chapter 7 (the final chapter) a summary of the main conclusions from the work presented in this thesis will be outlined. In addition, the chapter will close with possibilities for extending the models in this thesis as well as avenues for future research.

2 Systems modelling in biology

2.1 Introduction

The importance of the role that Systems Engineers have in biomedical research has increased considerably in recent years. This thesis describes how relatively straightforward, but very powerful, mathematical modelling techniques such as state-space modelling based on compartmental modelling (Section 2.2) can be used to describe biological processes in order to provide a practical and meaningful outcome. As a result this also allows further targeted research. The need for highly effective collaboration with biologists and mathematicians is demonstrated, where *real* data are generated to validate the mathematical models, and then predictions from models can be tested in the laboratory and clinic. This collaboration has blossomed to become a highly important niche, commonly termed as mathematical biology or for more complex systems, systems biology; mathematical biology is the application of mathematical modelling techniques to solve problems in biology and physiology [16]. Systems biology is the field of science that aims at system-level understanding of biology and exploring the interactions between biological systems and how these interactions give rise to functional biological properties of these systems [17].

A mathematical model (or model for short), irrespective of domain or application, is an abstract description of the behaviour of a system (in a real-world situation) in terms of mathematical equations [18]. A model of a particular system is constructed based on scientific (physical) laws that the elements of the system and their inter-connections (inter-relations) are believed to obey. The model type depends on the objective of the Systems Engineer (or Biological Modeller) as well as the tools for analysing the system. The tools that will be used in exploring the kinetic and dynamic models in this thesis are described in this chapter.

Building mathematical models involves imagination and skill. The following summary is an outline of the modelling process [19]:

- **Formulating the problem:** the nature of the selected model depends on what one seeks to find as well as the objective of developing it;
- **Outlining the model:** the parts of the system must be divided into unimportant (elements that have no effects and can be neglected), exogenous variables (also called parameters, inputs or independent variables) and endogenous (output or dependent) variables. In addition, the inter-relations (which constitute the assumptions of the model) between variables must be specified;
- **Verifying usefulness:** in this step, it is essential to check the possibility of obtaining the needed data that are then used in the model to make the expected predictions. If this is not possible, then the problem and/or the model should be reformulated.
- **Testing the model:** at this stage, the model is examined against data or hypotheses. However, relying on common sense alone is not advisable as it may well be wrong. If the predictions are not acceptable (with no mathematical errors) then adjustments should be made to steps one and/or two. If the predictions are less accurate than expected, then is it important to understand the reason behind that as there might be implicit or false assumptions.

The rationale behind biological systems modelling is to proceed with the quantitative understanding of biological functions as well as to improve medical science and practice by employing quantitative or qualitative knowledge [20]. In modelling biological processes, the objective is to summarise the available experimental evidence about the key (functional) characteristics of biological processes by deriving mathematical relations among variables of biological significance. The resulting mathematical models are expected to imitate the observed functional behaviour of the biological system.

In an ideal world, such a biological model should be compact (encompass the minimum computational and mathematical complexity), globally accurate (reproduce precisely the observed data from experiments under all natural operating conditions) and interpretable

(be consistent with biological investigation and analysis that advances the underlying mechanisms assisting the function of the biological process) [20]. Robustness of the biological model is another important attribute in modelling a system. That is, the mathematical model should provide stable behaviour against internal and/or external random perturbations (including noise). Therefore, the modelling approach can possibly be formed in a stochastic context to account for such perturbations. Developing a global biological model presumes the ability to measure (observe) the instantaneous activity of the variables of interest accurately with a high sampling resolution over time intervals.

However, in practice, the modelling process has been subject to restrictions in the ability to physically access functional characteristics of biological systems. In many engineering and scientific applications, the examined system can sometimes be studied and analysed entirely through invasive methods. In biological applications, invasive and non-invasive techniques can sometimes be employed to gather the required data; however, this is not always possible or ethical due to the delicate nature of such structures, particularly in humans. Therefore, one must rely on *in vitro* and animal studies which can sometimes be misleading if the physiology with respect to drug kinetics and dynamics is quite different compared to that in humans [21]. In addition to data acquisition problems, mathematical models describing biological systems are composed of sets of non-linear ordinary or partial differential equations which must be solved numerically by using digital computers. Advanced models that describe biological processes may require extensive computer services, and hence, the modelling process can be constrained due to the limited availability of computational (software packages) and analytical methods. This leads to the development of models that are “less than ideal” in the sense defined above. That is, an inability to achieve the desired features of global accuracy, compactness and interpretability.

Once a mathematical model of a biological process has been developed (built), it can be used in exploring and studying the response to various differing perturbations. One advantage of mathematical modelling is to infer from the model information relating to parts of the biological system that are not directly measured. Another advantage for using such models is the flexibility to perform simulations (using software packages) over different time intervals, this allows predictions about future (possibly long-term) behaviour

of the underlying biological system. Therefore, mathematical models can significantly reduce the resources (i.e., the cost) required to examine complex biological systems. This can be achieved by re-running the tests or multiple scenarios using powerful computing facilities, since repeating experiments using real-world facilities would incur considerable cost. On other occasions, recreating some experiments in the real-world is not possible due to multiple reasons including danger of recovery that might result from testing a drug on humans. Accordingly, mathematical modelling of biological processes *in vitro* to validate clinical methods is an important prerequisite to moving into further stages in drug trials to reduce the effects on living organisms to a minimum. Therefore, mathematical modelling has the potential to maximise the outcomes of biomedical and biological research. The methodology of mathematical modelling of systems is discussed comprehensively in [18, 19].

The models (Chapters 4-6) described in this thesis are based on underlying biological assumptions in which biochemical mechanisms are translated into systems of non-linear ordinary differential equations (ODEs) by the Law of Mass Action which describes the rate at which chemicals interact in reactions. Consider the following general chemical reaction:

$$aA + bB + \cdots + zZ \xrightarrow{k} P$$

where A, B, …, Z are chemical reactants to form a product chemical P, k is a proportionality constant known as a rate constant of the reaction, and a, b, …, z represent the number of moles of A, B, …, Z, respectively. The Law of Mass Action states that the rate of this chemical process is proportional to the product of the activities (or concentrations) of the reactants [22]. Mathematically, this is described by the following rate equation:

$$\frac{\mathrm{d}P}{\mathrm{dt}} = k[A]^a[B]^b \cdots [Z]^z .$$

The sum of the exponents, i.e., $(a+b+\cdots+z)$, in the rate equation determine the order of the chemical reaction. That is, the order corresponds to the molecularity[3] of the reaction. Accordingly, the chemical reaction $A \xrightarrow{k} P$ is an example of a first-order (or

[3] Molecularity: is the number of molecular entities that simultaneously collide in a chemical reaction.

unimolecular) reaction, whereas $2A \xrightarrow{k} P$ and $A + B \xrightarrow{k} P$ are examples of second-order (or bimolecular) reactions.

The ODEs are used to analyse the qualitative and quantitative behaviour of the underlying system to bridge the gap between mechanisms and theoretical biology. Numerical solution (using software packages) of the ODEs can then be compared with experimental data (observations) to determine unknown parameters of the models presented in this thesis.

The experimental data used in this thesis have been provided by the project collaborators (Wales College of Medicine, Cardiff University). The first data set which was used in the model development and parameter estimation of the drug kinetics model and the coupled drug kinetics/dynamics model (see Chapters 4 and 6) was collected from an *in vitro* study on the interaction of a culture of live human breast cancer cells (MCF-7 cell line) and TPT using two-photon laser-scanning microscopy (TPLSM). The second data set that was used in the model development and parameter estimation of the cell cycle model (see Chapter 5) and the coupled drug kinetics/dynamics model (see Chapter 6) was obtained from *in vitro* experiments on the cell cycle response using high-resolution Cyclin B1 fluorescence data from live human osteosarcoma cells (U-2 OS cell line) in the absence and presence of the anti-cancer agent TPT, respectively. The third data set that was used in the model development and parameter estimation of the coupled drug kinetics/dynamics model (see Chapter 6) was obtained from *in vitro* experiments (on U-2 OS cells) for the detection of H2AX phosphorylation foci (using high-content imaging) for determining the action of TPT that induces DNA damage (i.e., DNA damage data, see Chapter 6). Therefore, given the data available, the state-space modelling approach (based on compartmental modelling techniques) was used in this thesis to provide realistic mathematical models that describe biological processes, i.e., the *in vitro* kinetics and dynamics of the anti-cancer agent TPT.

State-space modelling techniques are applied extensively in biology using available computational and analytical mathematical tools to mathematicians and engineers. Applications extend from cell based (microscopic level) models up to the macroscopic level. Developed models are applied for predictive purposes, i.e., to assist in analysing and diagnosing treatment applications. Developing kinetic systems in addition to analysing and

simulating drug release into the body is essential in studying dynamic effects leading to drug discovery.

2.2 Compartmental modelling

A compartmental model (Figure 2.1) consists of a finite number of macroscopic subsystems (subunits) called compartments or pools, each of which is assumed to consist of homogenous and well-mixed material [23, 24]. The compartments interact by material flowing from one compartment to another and they can also exchange material with the environment. The basic law governing the flow into and out of compartments in such systems is the *Law of Mass Balance* which states that the matter can neither be created nor destroyed. That is, if quantity enters the system it either leaves or it accumulates within the system. The rate of change of quantities of materials within the compartments can be described by first-order ODEs. The most general form of compartmental equations for a (linear or nonlinear) system of n compartments is given by the mass balance equation:

$$\frac{\mathrm{d}x_i}{\mathrm{dt}} = f_{i0} + \sum_{\substack{j=1 \\ j \neq i}}^{n} f_{ij} - \sum_{\substack{j=1 \\ j \neq i}}^{n} f_{ji} - f_{0i} \tag{2.1}$$

where x_i (i = 1, …, n) is the quantity of material in compartment i, f_{ij} is the flow rate to compartment i from compartment j, and the environment is denoted by the subscript 0. A compartmental system can be used to model the kinetics of one substance, in which case the compartments occupy different physical spaces and the inter-compartment transfers represent the flow of material from one location to another. Moreover, such systems may also be used to model the kinetics of two or more substances, examples include a drug and its metabolites in which case different compartments may occupy the same physical space and some of the inter-compartment transfers represent the chemical transformation from one substance to another [25]. Compartmental models are applied extensively both in the physical and life sciences such as biomedicine/biomedical control (tumour targeting), chemical reaction systems (enzyme chains) in addition to pharmacokinetics and anaesthesia (drug kinetics).

State-space modelling falls under two broad headings: the *forward problem* and the *inverse problem* [23]. In the former, i.e. the forward problem, the structure of the system is known, that is the number of variables, their connectivities and the values of the non-zero transfer coefficients. Using this information, the behaviour of the system can be examined by simulating the state-space model. However, in real applications (experimental sciences), that is not the way state-space systems are generally applied. Therefore, the second approach, i.e. the inverse problem approach was applied in this thesis. That is, given some of the behaviour of a real system, infer the structure of the model and estimate the values of its unknown parameters using data from experiments.

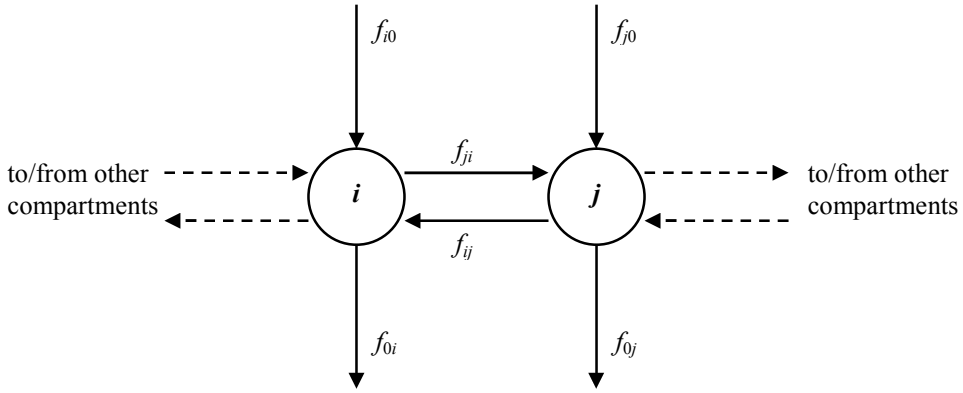


Figure 2.1: Two general compartments of a compartmental system, as described by Equation (2.1).

In biological/biomedical applications, systems models can be divided into main types based on the technique employed to construct the model; these models can be *models of data* or *models of process* or a mixture of both [26]. Models of data are chosen to fit particular data sets without considering the underlying mechanisms (basic processes) at work, i.e., there are no interpretations for the processes involved in material exchange between compartments in models of data. Examples of such models are linear and polynomial regression models of statistics [23]. The second type encompasses models of process, these models are built from models of the underlying mechanisms (or processes) at work in a biological system. Non-mechanistic models are less informative than mechanistic models since it is more difficult to address the issues of enhancement and optimisation in the former type of models (i.e., models of data). That is, in non-mechanistic models, there is no clear understanding for the compartments and/or the processes involved in exchanging material between the compartments. For these reasons, the state-space

models (see Chapters 4-6) derived in this thesis are models of processes based on the biological mechanisms explained in Chapter 3.

2.3 Applications of state-space models in biology

Many of the most common processes arising in biological and biomedical applications are non-linear and can be described by non-linear state-space models that are relatively complicated and require numerical simulation studies as analytical solutions are rarely available (see Section 2.1). Examples include the modelling of metabolic pathways, transport across membranes, pharmacokinetics and pharmacodynamics [23, 24]. In non-linear state-space models, the transfer rate coefficients between compartments (from compartment j to compartment i) can be a function of the amounts in all compartments, time, and a vector of rate constants $\mathbf{k}$. Therefore, based on Equation (2.1), the rate of change of the amount in compartment i (x_i ($i = 1, \ldots, n$)) is given by:

$$\frac{\mathrm{d}x_i}{\mathrm{dt}} = f_{i0} + \sum_{\substack{j=1 \\ j \neq i}}^{n} f_{ij}(\mathbf{x}(t), \mathbf{k}, t) x_j(t) - \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} f_{ji}(\mathbf{x}(t), \mathbf{k}, t) - f_{0i}(\mathbf{x}(t), \mathbf{k}, t) \right) x_i(t). \qquad (2.2)$$

A non-linear state-space model is *autonomous* if none of the transfer rate coefficients depend explicitly on time (i.e., of the form $f_{ij}(\mathbf{x}(t), \mathbf{k})$), otherwise the state-space model is *non-autonomous* (i.e., of the form $f_{ij}(\mathbf{x}(t), \mathbf{k}, t)$). Moreover, if the rate of transfer of material from compartment i to any other compartment or the environment depends only on the amount or concentration of that compartment (x_i) and that is true for all i, then the model is said to be a *donor-controlled* state-space model [23]. Otherwise, the state-space model is said to be a *donor-acceptor controlled* model. The state-space models (see Chapters 4-6) describing the biological processes (Chapter 3) in this thesis are autonomous, non-linear state-space models. Moreover, the kinetics and dynamics of TPT in addition to the cell cycle response involve enzymatic reactions as well as transfers across the cell membrane; the transfer rates in such processes depend on the concentration of other compartments. Therefore, the models described in this work are donor-acceptor controlled systems.

2.3.1 Enzyme kinetics

Enzyme mediated reactions are important biological processes that take place in cells and can be described mathematically using state-space models. The simplest form of enzymatic reaction is the irreversible conversion of substrate *s* into the product *p* by an enzyme *e* in a homogeneous reaction in a well-stirred reaction vessel of fixed volume. The standard form of an irreversible chemical reaction is given by [27]:

$$s + e \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} c \xrightarrow{k_2} p + e\,. \tag{2.3}$$

The substrate (*s*) combines with the free enzyme (*e*) to form an intermediate complex (*c*). The complex *c* can break down in two ways: it may release the free enzyme and substrate, or it may release the free enzyme and product. Let $S(t)$ be the concentration of the substrate, $P(t)$ be the concentration of the product, $C(t)$ is the concentration of the enzyme-substrate (intermediate) complex and $E(t)$ is the concentration of the free enzyme. The concentrations of $S(t)$, $P(t)$, $C(t)$ and $E(t)$ are measured in molar (M). The rate at which the substrate (*s*) binds to the enzyme (*e*) is proportional (with the association rate constant k_1) to the product of substrate and free enzyme $E(t)$. Dissociation of bound substrate (intermediate complex) occurs at a first-order rate as either substrate (*s*), with rate constant k_{-1} or product (*p*), with rate constant k_2. The rate constants k_1, k_{-1} and k_2 have the units of s^{-1} M^{-1}, s^{-1} and s^{-1}, respectively. Using Law of Mass Action, the equations for the rate of change of each concentration are:

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{dt}} &= -k_1 SE + k_{-1} C \\
\frac{\mathrm{d}E}{\mathrm{dt}} &= -k_1 SE + (k_{-1} + k_2) C \\
\frac{\mathrm{d}C}{\mathrm{dt}} &= k_1 SE - (k_{-1} + k_2) C \\
\frac{\mathrm{d}P}{\mathrm{dt}} &= k_2 C
\end{aligned}\,. \tag{2.4}$$

The state-space model describing the irreversible enzyme kinetics (in Equation (2.4)) is non-linear due to the product terms (2^{nd} order terms) in the concentrations. Note that,

$$\frac{\mathrm{d}E}{\mathrm{dt}} + \frac{\mathrm{d}C}{\mathrm{dt}} = 0, \tag{2.5}$$

that is, $E(t) + C(t)$ is constant for a closed system. In addition,

$$\frac{\mathrm{d}S}{\mathrm{dt}} + \frac{\mathrm{d}C}{\mathrm{dt}} + \frac{\mathrm{d}P}{\mathrm{dt}} = 0, \tag{2.6}$$

that is, $S(t) + P(t) + C(t)$ is also constant for a closed system. Chemically, the enzyme appears in two forms, as free enzyme (e) or as the intermediate complex (c). The complex c plays a double role as far as the conservation of substance is concerned [23]; that of the substrate (s) in the form of c and that of the enzyme (e) in the form c.

The initial conditions for the model are:

$$S(0) = S_0,\ E(0) = E_0,\ C(0) = 0 \text{ and } P(0) = 0. \tag{2.7}$$

If E_0 is the total concentration of available (free) enzyme, therefore, integrating Equation (2.5) with respect to time t subject to Equation (2.7), the concentration of $E(t)$ is the difference between E_0 and the intermediate complex (bound substrate), that is:

$$E(t) = E_0 - C(t). \tag{2.8}$$

Taking the rates of change of $S(t)$, $C(t)$ and $P(t)$ in Equation (2.4) and substituting from Equation (2.8) gives:

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{dt}} &= -k_1 S E_0 + (k_1 S + k_{-1})C \\
\frac{\mathrm{d}C}{\mathrm{dt}} &= k_1 S E_0 - (k_1 S + k_{-1} + k_2)C\,. \\
\frac{\mathrm{d}P}{\mathrm{dt}} &= k_2 C
\end{aligned} \tag{2.9}$$

Equation (2.9) represents the new state-space model describing the irreversible conversion of the substrate s to the product p where $E(t)$ has been eliminated to reduce the number of state variables. This technique has been applied in the extended drug kinetics model (see Chapter 4) to mathematically describe the irreversible conversion of the active form of the drug to the inactive form which is catalysed by the enzyme aldehyde dehydrogenase (ALDH, see Chapter 3).

2.3.2 Pseudo-steady state approximation

The purpose of this subsection is to derive Michaelis-Menten kinetics [28]. Pseudo-steady state approximation (PSSA) is a standard approximation that is widely applied to models in pharmacokinetics, biology and biomedicine. This technique is based on the following assumptions (using the state-space model of enzyme kinetics in Section 2.3.1 as the exemplar) [27]: enzyme (e) is present in a relatively small concentration in comparison to the substrate (s), that is, $E(t) << S(t)$ and/or the binding affinity of the enzyme is much lower than unity, that is, $(k_{-1} + k_2)/k_1 << 1$. According to these assumptions, the intermediate compound (enzyme-substrate complex c) reaches an equilibrium state very rapidly, i.e., $\frac{\mathrm{d}C}{\mathrm{dt}} = 0$, very shortly after the start of the reaction. Solving for C in Equation (2.9) results in:

$$C = \frac{k_1 S E_0}{k_1 S + k_{-1} + k_2}. \tag{2.10}$$

Using the relation in Equation (2.10) and substituting in the state-space model for irreversible enzymatic reactions, Equation (2.9) can be reformulated and simplified to Michaelis-Menten approximation:

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{dt}} &= \frac{-V_m S}{S + k_m} \\ \frac{\mathrm{d}P}{\mathrm{dt}} &= \frac{V_m S}{S + k_m} \end{aligned} \tag{2.11}$$

where $k_m = (k_{-1} + k_2)/k_1$ and $V_m = k_2 E_0$. In Equation (2.11) V_m is the relevant maximum reaction rate and k_m is the relevant Michaelis-Menten constant at which the reaction rate is one-half of V_m [23]. The PSSA presented in this subsection will be used to reduce the number of unknown parameters in the extended drug kinetics model (see Chapter 4) by applying Michaelis-Menten Law to the catalytic reaction of the enzyme ALDH (see Chapters 3 and 4) and to the efflux pumping mechanism via the breast cancer resistance protein (BCRP)/ABCG2 drug transporter (see Chapters 3 and 4).

2.3.3 Transport across membranes

The cell membrane (plasma membrane) is a lipid barrier that physically separates the interior of all cells and the environment (exterior of cells). Due to the lipid nature of the plasma membrane it is highly impermeable to both charged[4] and polar[5] solutes, therefore, this biological membrane is able to regulate what enters and exits the cell. However, important solutes are either charged (such as amino acids as well as ions) or polar (such as glucose). There are several types of transport mechanisms across the cell membrane, that is, cells have developed highly selective mechanisms that facilitate either the entry or exit of important metabolites, by active transport (against the concentration gradient) by moving them up free energy gradients [23]. All of the selective mechanisms utilise proteins that are in the cell membrane which have special binding sites for the transported (solutes) molecules. Once the solute is bound to the binding site of its transporter (carrier), the complex can change conformation so that the bound solute can be discharged (released) on the other side of the membrane. The transporting mechanism (Figure 2.2) that is modelled in this subsection is a simplified active pumping mechanism for transporting a solute across a cell membrane. Additionally, in this thesis, it is assumed that the transported solute is not metabolised in the cell membrane.

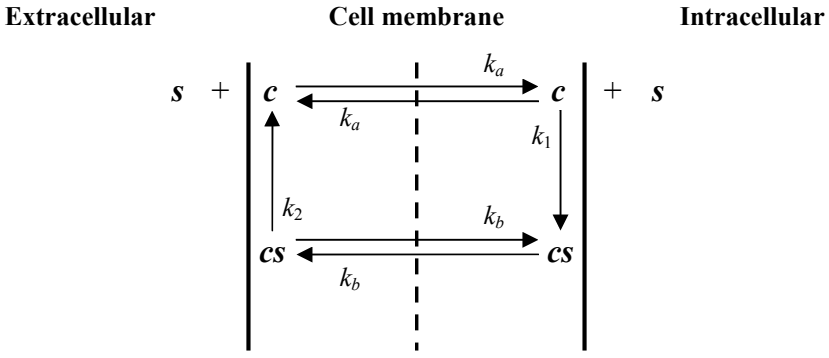


Figure 2.2: Schematic of symmetric carrier model for active transport.

The membrane contains a carrier (transporter), *c*, that can combine with a specific solute (substrate), *s*, to form a transporter-substrate complex (intermediate complex) *cs*. Both, the

[4] Charged solute: having an electric charge

[5] Polar solute: a solute in which there is some separation of charge in the chemical bonds. Hence, one part of the molecule has a slight positive charge and the other has a slight negative charge. That is, the solute has a permanent electric dipole.

free transporter, *c*, and the transporter-substrate complex, *cs*, exist in two states: one in which the binding reaction is with the substrate in the intracellular region, the other in which the transporter dissociates the substrate to the extracellular region. For the transporter (*c*) and the transporter-substrate complex (*cs*), there are transition rates between the two states (see Figure 2.2). Similar to the approach used by Jacquez [23], it is assumed that the model in Figure 2.2 is a symmetric active transporter model, i.e., the rate constants for the transition between the two states for the transporter are equal (denoted by k_a) and rate constants for the transition between the two states of the transporter-substrate complex are equal (denoted by k_b). The intermediate complex (*cs*) breaks down in one direction to release the substrate at the extracellular region (see Figure 2.2).

Let $S_1(t)$ be the concentration of the substrate in the intracellular region, $S_2(t)$ be the concentration the substrate in the extracellular region, $CS(t)$ is concentration of the transporter-substrate complex and $C(t)$ is the concentration of the free transporter. The rate at which the substrate (*s*) binds to the transporter (*c*) is proportional (with association rate constant k_1) to the product of substrate and free transporter $C(t)$. Dissociation of bound substrate (intermediate complex) occurs at a first-order rate as substrate (*s*) in the extracellular region, with rate constant k_2.

If C_0 denotes the total concentration of available transporter (see Subsection 2.3.1), then the concentration of free transporter is the difference between C_0 and the concentration of the transporter-substrate complex (CS), that is $C(t) = C_0 - CS(t)$. If the volumes of the extracellular region, cell membrane and intracellular region are denoted by V_a, V_b and V_c, respectively, the mathematical model for the active transport mechanism (described in this subsection) is given by the following system of differential equations given that $v_1 = V_b/V_c$ and $v_2 = V_a/V_b$.

$$\begin{aligned} \frac{\mathrm{d}S_1}{\mathrm{dt}} &= -k_1 S_1 (C_0 - CS) \\ \frac{\mathrm{d}CS}{\mathrm{dt}} &= \frac{k_1 S_1 (C_0 - CS)}{v_1} - k_2 CS \,. \\ \frac{\mathrm{d}S_2}{\mathrm{dt}} &= \frac{k_2 CS}{v_2} \end{aligned} \tag{2.12}$$

Assuming that the concentration of the transporter is much lower than the concentration of associated substrate (i.e., the transporter is saturable), PSSA can be applied to the model described by Equation (2.12), see Subsection 2.3.2. The transporter (*c*) is present in a relatively small concentration in comparison to the substrate (*s*) in the intracellular region, that is, $CS(t) << S_1(t)$ and/or the binding affinity of the transporter is much lower than unity, that is, $k_2/k_1 << 1$. According to these assumptions, the intermediate compound (transporter-substrate complex *cs*) reaches an equilibrium state very rapidly, i.e., $\frac{\mathrm{d}CS}{\mathrm{dt}} = 0$, very shortly after the start of the active transport. Solving for *CS* in Equation (2.12) results in:

$$CS = \frac{k_1 S_1 C_0}{k_1 S_1 + k_2 v_1}. \tag{2.13}$$

Using the relation in Equation (2.13) and substituting in the state-space model (Equation (2.12)), the equations for S_1 and S_2 can be reformulated and simplified to Michaelis-Menten approximation (see Subsection 2.3.2):

$$\begin{aligned} \frac{\mathrm{d}S_1}{\mathrm{dt}} &= \frac{-V_m S_1 v_1}{S_1 + k_m v_1} \\ \frac{\mathrm{d}S_2}{\mathrm{dt}} &= \frac{V_m S_1}{v_2 (S_1 + k_m v_1)} \end{aligned} \tag{2.14}$$

where $k_m = k_2/k_1$ and $V_m = k_2 C_0$. In Equation (2.14) V_m is the relevant maximum reaction rate and k_m is the relevant Michaelis-Menten constant at which the reaction rate is one-half of V_m [23]. This technique is applied in the extended drug kinetics model (see Chapter 4) to mathematically describe active pumping via the drug transporter.

2.3.4 The relationship between response and drug concentration

Wagner [29] proposed that the intensity of a pharmacologic response (i.e. the pharmacological effect), *R* (at time *t*), is related to the concentration of a drug (in the vicinity of the receptor), *C* (at time *t*), by the following Hill equation:

$$\frac{R}{R_{\max}} = \frac{C^n}{K + C^n} \tag{2.15}$$

where R_{max} is the maximum intensity of the pharmacologic response possible; $R \rightarrow R_{max}$ as $C \rightarrow \infty$. K is a constant and the superscript n is a positive constant. A plot of the function R versus log$C$ produces a sigmoidal shape which is approximately linear over the range $0.2R_{max} < R < 0.8R_{max}$ and therefore, the relationship between the intensity of a pharmacologic response ($R(t)$) and the drug concentration ($C(t)$) becomes [24]:

$$R = m \log C + e \tag{2.16}$$

where m and e are constants. The relationship in Equation (2.16) between $R(t)$ and $C(t)$ is normally suitable over a wide proportion of the therapeutic range. In order to obtain the response-time curve, Equation (2.16) (or Equation (2.15) if a wide therapeutic range is covered) is applied to the drug concentration-time curve ($C(t)$). As the concentration $C(t)$ becomes closer to a single exponential, the more nearly the intensity of the pharmacologic response $R(t)$ approximates to a straight line. Hence, for a one-compartment model, $R(t)$ is a straight line over a large proportion of the therapeutic range (for a single rapid intravascular drug administration) and approximates a straight line for relatively large values of t (for extra-vascular drug administration).

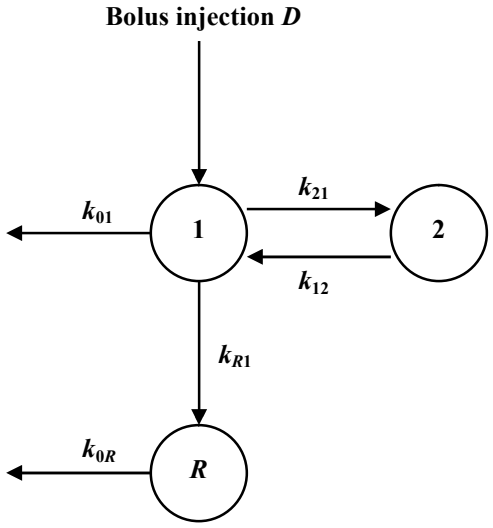


Figure 2.3: Two-compartment drug kinetic model with response (effect) compartment (*R*) to model the pharmacological response (redrawn from Reference [24]).

Finding the appropriate curve for $C(t)$, in multi-compartment drug kinetics models, is generally not an easy task. For example, taking an n-compartment model ($n > 1$) it is often found that the substitution of either $C_1(t)$, …, or $C_n(t)$ ($i = 1, \ldots, n$) into Equations (2.15) or (2.16) gives a poor estimation of the pharmacologic response-time curve $R(t)$ [24]. This

problem has been overcome by two methods: The first method is to extend the drug kinetics model by introducing a response (effect) compartment denoted by the variable $R(t)$, as shown in Figure 2.3. The idea of adding a response (drug effect) compartment was introduced to the field of pharmacokinetics by Sheiner *et al.* [30] and to the field of anaesthetics by Hull *et al.* [31]. The response compartment R has a *negligible* effect on the drug kinetics model, that is, the rate constant k_{R1} is relatively small in comparison to the other rate constants, namely, k_{01}, k_{12} and k_{21}, i.e., $x_1(t)$ is unchanged [24]. Accordingly, if compartments 1 and 2 are denoted by the variables $x_1(t)$ and $x_2(t)$, respectively, then the rate of change of the quantity in the response compartment is given by (using Equation (2.1)):

$$\frac{\mathrm{d}R}{\mathrm{dt}} = -k_{0R}R + k_{R1}x_1 \,. \tag{2.17}$$

In this particular example, the order of the response (effect) compartment has increased by one (the response compartment is third-order) and, according to Godfrey [24], enables the dynamic data (response data) to be fitted more closely in comparison to the approach described in Equations (2.15) and (2.16). The second method is to model the response as a sum of contributions from the compartments of the drug kinetics model (see the text by Godfrey [24] and the paper by Whiting and Kelman [32]); this method is not within the scope of this thesis. Modelling the dynamic effect of the anti-cancer agent on live human osteosarcoma cells is described in Chapter 6 using the first method in Equation (2.17).

2.4 Identifiability of model parameters

2.4.1 Structural identifiability

The major techniques for the inverse problem (see Section 2.2) involve the determination of the structure of the underlying model and parameter estimation. An important objective is to estimate the unknown parameters of the postulated system and often it is difficult to observe all of the state-space variables (compartments) of the state-space model. In

possible experiments (especially for biological systems) one is often restricted so that only a small subset of the system can be observed. When developing a parametric state-space model of a given physical system, such as drug kinetics, the basic question of whether the parameters can be identified (uniquely or otherwise) from a specified input-output experiment (Figure 2.4) arises. The question concerns whether there is enough information in the observations (measurements) to uniquely determine the unknown parameters of the system. Structural identifiability analysis (see, for example, [33] and the papers in [34]) can be applied to try to resolve this question in the theoretical situation of perfect, noise-free and continuous observations. This is a question relating to the structure of the state-space model, and deals with whether different combinations of parameters give rise to the same (ideal) measurement data (observations) [35]. Moreover, this problem is independent of the actual experimental data (measurements) collected. It should be noted that structural identifiability is an *a priori* analysis, i.e. it can be performed before obtaining experimental data (measurements). However, in practice, data are noisy and never perfect and therefore, it is essential to analyse the effect of errors in experimental data (observations). Hence, the question of identifiability in such data is known as the *numerical* or *a posteriori* problem (see Section 2.5).

Figure 2.4: Input-output system.

Structural identifiability analysis is an important step in the modelling process and the output structure has to uniquely determine the unknown parameters in a postulated model in order to approach the estimation of the unknown model parameters with greater confidence. If the unknown parameters are unidentifiable by the input-output behaviour, the proposed experiment should be redesigned (if possible) in order to obtain an output structure that uniquely determines the unknown parameters (globally or locally). This would involve changing which functions of the states are observed (the observations/output structure) and/or changes can also include how input(s) enter the system (the input structure). In some cases, it is not possible to achieve an identifiable model after redesigning the input-output structure and that is due to the limitation of the process being modelled [36]. Therefore, any subsequent estimation of the unknown model

parameters should be viewed with less confidence. The main objective of identification problems in industry and biosciences is to combine *a priori* knowledge with experimental data [33]. Accordingly, a structural identifiability analysis is a prerequisite for experiment design, system identification and parameter estimation.

To formally define structural identifiability in mathematical terms, it is necessary to introduce a general class of parametric models for which the state-space models derived in this thesis are particular examples. Consider a postulated parametric model given by:

$$\begin{aligned} \dot{\mathbf{x}}(t,\mathbf{p}) &= \mathbf{f}(\mathbf{x}(t,\mathbf{p}),\mathbf{p}) \\ \mathbf{x}(0,\mathbf{p}) &= \mathbf{x}_0(\mathbf{p}) \\ \mathbf{y}(t,\mathbf{p}) &= \mathbf{h}(\mathbf{x}(t,\mathbf{p}),\mathbf{p}) \end{aligned} \tag{2.18}$$

where $\mathbf{p} = (p_1, \ldots, p_q)^{\mathrm{T}}$ is a q-dimensional vector representing the unknown model parameters and lies in some open subset $\Omega \subset \mathbb{R}^q$ of feasible vectors. The n-dimensional vector $\mathbf{x}(t, \mathbf{p}) = (x_1(t, \mathbf{p}), \ldots, x_n(t, \mathbf{p}))^{\mathrm{T}}$ is the state vector and represents the information required by the system to evolve in time, where $x_i(t, \mathbf{p})$ are the model variables. Solutions to Equation (2.18) are assumed to lie in $M(\mathbf{p})$ which is the largest connected open subset of $\mathbb{R}^n$ (i.e., all n components are real numbers) such that $\mathbf{x}(t, \mathbf{p}) \in M(\mathbf{p})$ for all $t \geq 0$. It is assumed that both $\mathbf{f}(\cdot, \mathbf{p})$ and $\mathbf{h}(\cdot, \mathbf{p})$ are rational functions in both $\mathbf{x}$ and $\mathbf{p}$. The vector $\mathbf{y}(t, \mathbf{p}) = (y_1(t, \mathbf{p}), \ldots, y_r(t, \mathbf{p}))^{\mathrm{T}}$ is the response function or the observation function (functions of the state that are measured in experiments), where $\mathbf{y}(t, \mathbf{p}) \in \mathbb{R}^r$. It is also assumed that the initial condition $\mathbf{x}_0(\mathbf{p})$ is a rational function of $\mathbf{p}$.

The inputs in the experiments considered in this thesis consist of a single impulse, i.e. injection of the anti-cancer drug TPT into the system, therefore, the corresponding amounts, are included in the initial conditions $\mathbf{x}_0(\mathbf{p})$. Otherwise, there are no external inputs to the system and no further input terms are included in Equation (2.18). According to the proposed experiments (see Chapter 4), r linear combinations of the variables of the model are measured, $c_{i1}(\mathbf{p})x_1(t, \mathbf{p}) + \cdots + c_{in}(\mathbf{p})x_n(t, \mathbf{p})$, $(i = 1, \ldots, r)$. Thus the following equation represents the output structure of the model for the systems considered in this thesis:

$$\mathbf{y}(t,\mathbf{p}) = \mathbf{C}(\mathbf{p})\mathbf{x}(t,\mathbf{p}) \quad \text{where} \quad \mathbf{C}(\mathbf{p}) = \begin{pmatrix} c_{11}(\mathbf{p}) & \cdots & c_{1n}(\mathbf{p}) \\ \vdots & & \vdots \\ c_{r1}(\mathbf{p}) & \cdots & c_{rn}(\mathbf{p}) \end{pmatrix}. \tag{2.19}$$

Definition 2.1: If a parameter (p_i) can take any of a distinct (countable) number of values for a given observation function, $\mathbf{y}(t, \mathbf{p})$, then the parameter ($p_i$) is *locally identifiable* (*a priori*).

Definition 2.2: If the parameter (p_i) can only take a unique value (in the entire parameter space) for the given observation function, then the parameter (p_i) is *globally/uniquely identifiable* (*a priori*) for that particular experiment.

Definition 2.3: Otherwise, the parameter (p_i) is unidentifiable (*a priori*).

Definition 2.4: A state-space model of the form in Equations (2.18) is said to be *structurally globally/uniquely identifiable* if all the unknown parameters of the model are globally/uniquely identifiable.

Definition 2.5: A state-space model of the form in Equations (2.18) is said to be *structurally locally identifiable* if all the unknown parameters of the model are locally identifiable and at least one parameter (p_i) is not globally identifiable.

Definition 2.6: Otherwise, a model of the form in Equations (2.18) is said to be *unidentifiable* if at least one of the unknown parameters (p_i) of the model is unidentifiable.

The definitions above are all based on the uncontrolled structure including the form of the input, the form of the input-output experiments as well as all the information about the parameter vector (**p**), others also consider classes of inputs [25]. These definitions apply to linear and non-linear state-space models in general.

The structural identifiability of the parameters for analysing linear, time-invariant systems expressed in state-space form is a well-established concept. There are many techniques discussed in Walter [34], particularly in the tutorial paper published by Godfrey and DiStefano III [25]. Among the most commonly used structural identifiability techniques are those based on the uniqueness properties of the Laplace transform (transfer function

approach) of the output (or outputs) [23, 25] or on the existence of a similarity transformation [15, 23-25]. In this thesis, the Laplace transform approach will be explained in detail (see Section 2.4.2). However, for non-linear state-space models, relatively few techniques are available, the similarity transformation technique [37, 38], differential algebra techniques [39, 40] and the most commonly used is the Taylor series approach [41]. The Taylor series approach (see Section 2.4.3), which will be used in this thesis, can be applied to state-space models where the drug administration is via a bolus injection (impulsive input) and performed by expanding the output (observations/measurements) as a Taylor series around time $t = 0$ [41]. This technique, i.e. Taylor series approach, can however be computationally intractable [24]. For non-linear models, significant problems still need to be resolved, for example, a problem that results from the consideration that global identifiability be sought with respect to a specific input [42] (including bolus injection input) of drug in pharmacokinetics.

2.4.2 The Laplace transform approach

In this subsection, linear compartmental systems with constant coefficients are studied to analyse the two-compartment model developed by Evans *et al.* [15] (see Section 4.2). Jacquez [23] considers the Laplace transform (or the transfer function) method as the easiest approach to use. In addition, this technique, is the most readily comprehended which immediately specifies the observational parameters (parameters that are incorporated in the observations) as functions of the basic parameters [23]. Bellman and Åström [33] indicated that for linear compartmental systems, if a model is identifiable from any input, the system is also identifiable from the impulse response. Accordingly, one can examine the transfer function for impulsive inputs at $t = 0$ for a compartmental system that starts with zero initial conditions (nothing in the compartments).

Consider a linear deterministic system with constant coefficients, zero initial conditions, and state-space representation of the following vector-matrix form:

$$\begin{aligned} \dot{\mathbf{x}}(t,\mathbf{p}) &= \mathbf{A}(\mathbf{p})\mathbf{x}(t,\mathbf{p}) + \mathbf{B}(\mathbf{p})\mathbf{u}(t);\ \mathbf{x}(0,\mathbf{p}) = 0 \\ \mathbf{y}(t,\mathbf{p}) &= \mathbf{C}(\mathbf{p})\mathbf{x}(t,\mathbf{p}) \end{aligned} . \tag{2.20}$$

The above is a state-space system of n first-order linear ODEs, similarly, $\mathbf{p} = (p_1, \ldots, p_q)^T$ is a constant q-dimensional vector representing the unknown (constant) model parameters and the n-dimensional vector $\mathbf{x}(t, \mathbf{p}) = (x_1(t, \mathbf{p}), \ldots, x_n(t, \mathbf{p}))^T$ is the state vector (n state variables of the state-space model). The vector $\mathbf{u}(t) = (u_1(t), \ldots, u_p(t))^T$ is the input (or control) vector, where $\mathbf{u}(t) \in \mathbb{R}^p$. The vector $\mathbf{y}(t, \mathbf{p}) = (y_1(t, \mathbf{p}), \ldots, y_r(t, \mathbf{p}))^T$ is the vector of r known output (observations) variables of the model . In Equation (2.20) $\mathbf{A}(\mathbf{p})$ is the matrix of transfer coefficients which includes the structural (basic) parameters of the model, $\mathbf{B}(\mathbf{p})$ is the input matrix and $\mathbf{C}(\mathbf{p})$ is the observational matrix. The matrices $\mathbf{A}(\mathbf{p})$, $\mathbf{B}(\mathbf{p})$ and $\mathbf{C}(\mathbf{p})$ are constants, each consisting of some or all of the unknown parameters and their dimensions are $(n \times n)$, $(n \times p)$, and $(r \times n)$, respectively.

Parameters can be classified into observational parameters and basic parameters [23, 43, 44]. The observational parameters are determined by the experimental design and are functions of a basic parameter set [23]. The basic parameter set consists of the kinetic parameters of the system (system invariants) in addition to other parameters that may also be introduced by an experiment.

The Laplace transform [18] provides a means of transforming $f(t)$ from the time domain to the complex-frequency domain. The defining equation for the Laplace transform of $f(t)$, i.e., $F(s)$, is given by:

$$F(s) = \int_0^\infty f(t)e^{-st}dt$$

where the complex variable s is written as:

$$s = \sigma + j\omega.$$

Taking Laplace transform of Equations (2.20) and rearranging equations for further analysis of the identifiability properties of linear constant coefficients models gives:

$$s\mathbf{X}(s) = \mathbf{A}(\mathbf{p})\mathbf{X}(s) + \mathbf{B}(\mathbf{p})\mathbf{U}(s) \tag{2.21}$$

and rearranging for $\mathbf{X}(s)$ results in:

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A}(\mathbf{p}))^{-1}\mathbf{B}(\mathbf{p})\mathbf{U}(s). \tag{2.22}$$

The Laplace transform of the observation function is:

$$\mathbf{Y}(s) = \mathbf{C}(\mathbf{p})(s\mathbf{I} - \mathbf{A}(\mathbf{p}))^{-1}\mathbf{B}(\mathbf{p})\mathbf{U}(s). \tag{2.23}$$

Recall that $\mathbf{C}(\mathbf{p})(s\mathbf{I} - \mathbf{A}(\mathbf{p}))^{-1}\mathbf{B}(\mathbf{p})$ is the transfer function for the model of experiment. If $\mathbf{D}(s) = adj(s\mathbf{I} - \mathbf{A}(\mathbf{p}))$, where adj(M) is the adjoint of matrix M, then,

$$(s\mathbf{I} - \mathbf{A}(\mathbf{p}))^{-1} = \frac{\mathbf{D}(s)}{|s\mathbf{I} - \mathbf{A}(\mathbf{p})|} = \frac{\mathbf{D}(s)}{s^n + \phi_1 s^{n-1} + \ldots + \phi_n},$$

where ϕ_i, $i = 1, \ldots, n$ are the observational parameters and the transfer function $\mathbf{G}$(s) can be written as,

$$\mathbf{G}(s) = \frac{\mathbf{C}(\mathbf{p})\mathbf{D}(s)\mathbf{B}(\mathbf{p})}{s^n + \phi_1 s^{n-1} + \ldots + \phi_n}. \tag{2.24}$$

The next step is to examine the observation (output) function given by:

$$\mathbf{Y}(s) = \mathbf{G}(s)\mathbf{U}(\mathbf{p}) = \frac{\mathbf{C}(\mathbf{p})\mathbf{D}(s)\mathbf{B}(\mathbf{p})\mathbf{U}(\mathbf{p})}{s^n + \phi_1 s^{n-1} + \ldots + \phi_n}. \tag{2.25}$$

For each of the observation function components ($Y_i(s)$), the numerator is also a polynomial in s but of degree less than or equal to $n - 1$. Writing the numerators and denominators in factored form can possibly result in observation functions (i.e. $\mathbf{Y}(s)$) that have common factors in numerator and denominator which should be cancelled before writing $\mathbf{Y}(s)$ in the following standard form:

$$\mathbf{Y}(s) = \frac{\phi_{n+1} s^{n-1} + \ldots + \phi_{2n}}{s^n + \phi_1 s^{n-1} + \ldots + \phi_n}. \tag{2.26}$$

The polynomial in the denominator will be of degree n (the coefficient of s^n is 1 in the denominator) and the polynomial in the numerator will be of a lower degree except for the possibility of cancellation of common factors in numerators and denominators [23]. If some of the coefficients are not independent, the determinant of the Jacobian matrix ($\partial\phi_i / \partial\theta_j$, where θ_j, $j = 1, \ldots$ are the basic parameters) will be zero. The observational parameters (the coefficients, which are often called *moment invariants*) that appear in the numerator and denominator of $\mathbf{Y}(s)$ have a particular structure of interest. That is, each observational parameter is homogeneous in the basic parameters. The highest degree coefficient (ϕ_n) is of degree n in the basic parameters, in addition, ϕ_n is non-zero unless

there is a zero eigenvalue. There may be more than one ϕ of each degree in the numerator and the denominator except for that of the highest degree *n*, i.e., there is only one ϕ_n. Moreover, the lowest degree coefficients in both, the numerator and denominator are linear in the basic parameters that appear in them. Therefore, Equations (2.27) represent the set of equations for the resulting moment invariants (observational parameters) associated with Equation (2.26)

$$\begin{aligned}
\phi_1 &= \sum_i \theta_i \\
\phi_2 &= \sum_{i \neq j} \theta_i \theta_j \\
\phi_3 &= \sum_{i \neq j \neq k} \theta_i \theta_j \theta_k \\
&\vdots
\end{aligned} \tag{2.27}$$

The identifiability properties are established by examining the information available about the vector of unknown model parameters (**p**) as well as the expressions in the powers of *s* in the numerators and denominators of **Y**(*s*). The Laplace transform method presented in this section will be applied to perform the structural identifiability analysis for the two-compartment hydrolysis model in Chapter 4.

The model in Figure 2.5 is used to illustrate structural identifiability analysis using the Laplace transform approach. The following are the state-space equations for that model with **x**(0, **p**) = 0:

$$\begin{aligned}
\dot{\mathbf{x}}(t,\ \mathbf{p}) &= \begin{bmatrix} -a_{01} & a_{12} \\ 0 & -a_{12} \end{bmatrix} \begin{bmatrix} x_1(t,\ \mathbf{p}) \\ x_2(t,\ \mathbf{p}) \end{bmatrix} + \begin{bmatrix} 0 \\ b \end{bmatrix} u(t) \\
y(t,\ \mathbf{p}) &= \begin{bmatrix} c & 0 \end{bmatrix} \begin{bmatrix} x_1(t,\ \mathbf{p}) \\ x_2(t,\ \mathbf{p}) \end{bmatrix}
\end{aligned}$$

Using Equation (2.23) and rearranging gives:

$$\begin{aligned}
Y(s) &= \begin{bmatrix} c & 0 \end{bmatrix} \begin{bmatrix} s + a_{01} & -a_{12} \\ 0 & s + a_{12} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ b \end{bmatrix} \\
&= \frac{bca_{12}}{s^2 + (a_{01} + a_{12})s + a_{01}a_{12}}
\end{aligned}$$

where $Y(s)$ is the Laplace transform of $y(t,\ \mathbf{p})$. The coefficients are the observational parameters. Therefore, determining $Y(s)$ means that the values for all coefficients are

determined. The equations for the observational parameters in terms of the basic parameters are:

$$\begin{aligned}\phi_1 &= a_{01} + a_{12} \\ \phi_2 &= a_{01}a_{12} \\ \phi_3 &= bca_{12}\end{aligned} \quad .$$

The observational parameters ϕ_1, ϕ_2 and ϕ_3 are unique; ϕ_1 and ϕ_2 yield two possible solutions for a_{01} and a_{12}. In addition, if the basic parameter b (or c) is known, using ϕ_3 yields two possible solutions for c (or b) and hence, the model is structurally locally identifiable. If both basic parameters b and c are known, from ϕ_3, a_{12} is determined and finally from ϕ_1 or ϕ_2, a_{01} is determined and therefore, the model is structurally globally identifiable. If neither b nor c are known, the model is unidentifiable.

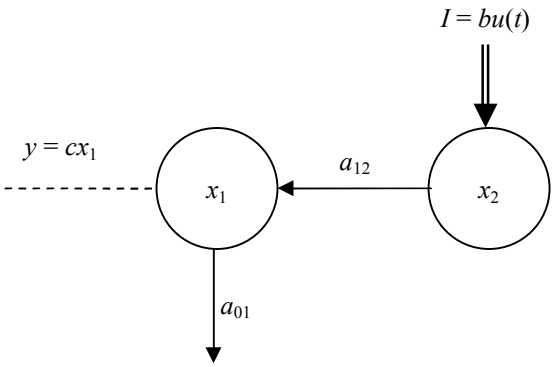


Figure 2.5: Two-compartment model to illustrate the Laplace transform approach.

2.4.3 Taylor series expansion of the observations

As mentioned earlier in this chapter, there are relatively few techniques available for testing the structural identifiability of non-linear state-space models expressed in the form of Equations (2.18) and (2.19). Deterministic identifiability for non-linear models presents even more of a problem than it did for linear models. For example, the superposition property no longer holds nor do relationships such as the step response being the integral of the impulse response [24]. Therefore, the transfer function (Laplace transform) method presented in the previous subsection no longer works for non-linear systems. However, the output structure (observation) is still available and contains all information about the

parameters of the state-space model that is available from the experiment. Thus an alternative approach of processing the observation is needed. The method that will be used in this thesis for testing the structural identifiability of non-linear state-space models (Sections 4.4, 5.4 and 6.4) is the Taylor series approach of Pohjanpalo [41]. This method has been applied extensively in models and experiments of interest in pharmacokinetics [45] and pharmacodynamics.

The Taylor series approach is in particular applicable to systems for which the administration is through bolus injection (impulsive input). An important condition to perform this test is that the observation must be analytic in a neighbourhood of $t = 0$, that is, the observation function $\mathbf{y}(t, \mathbf{p})$ is infinitely differentiable with respect to time t on this neighbourhood [42]. The Taylor series approach involves expanding the components of $\mathbf{y}(t, \mathbf{p})$, the observation function, and its successive derivatives are evaluated in terms of the model parameters and initial conditions $\mathbf{x}_0(\mathbf{p})$ as Taylor series around $t = 0^+$. For the system described in Equations (2.18) and (2.19), the Taylor series approach is used to verify that the system is structurally globally identifiable. Mathematically, the Taylor series expansion of the components of the observation function, $\mathbf{y}(t, \mathbf{p})$, about $t = 0^+$ is given by:

$$y_i(t,\ \mathbf{p}) = y_i(0,\ \mathbf{p}) + t y_i^{(1)}(0,\ \mathbf{p}) + \frac{t^2}{2!} y_i^{(2)}(0,\ \mathbf{p}) + \cdots + \frac{t^k}{k!} y_i^{(k)}(0,\ \mathbf{p}) + \cdots \ \ i = 1, \cdots, r$$

where

$$y_i^{(k)}(0,\ \mathbf{p}) = \lim_{\tau \to 0^+} \frac{\mathrm{d}^k y_i}{\mathrm{d}t^k}(\tau,\ \mathbf{p}),\ \ i = 1, \cdots, r.$$

The coefficients in the Taylor series expansion are unique (identifiable/measurable) functions of the basic parameters and therefore are observational parameters (i.e., they contain information about the unknown model parameters) [23, 42]. That is, uniqueness of the coefficients in the Taylor series expansions of $y_i(t, \mathbf{p})$ entails the following, if $\bar{\mathbf{p}} \in \boldsymbol{\Omega}$ is such that $\mathbf{y}(t, \mathbf{p}) = \mathbf{y}(t, \bar{\mathbf{p}})$ for all $t \geq 0$, then, for each $i = 1, \ldots, r$ and $k = 1, 2, 3, \ldots$

$$y_i^{(k)}(0, \mathbf{p}) = y_i^{(k)}(0, \bar{\mathbf{p}}) \ \text{ where, } y_i^{(0)}(0, \mathbf{p}) = y_i(0, \mathbf{p}). \tag{2.28}$$

In applying the Taylor series approach to non-linear systems there is usually an infinite set of coefficients, that is, the upper bound on the number of equations required for this set is

not easily determined. However, a certain number of equations are independent. For linear systems, by applying the Cayley-Hamilton theorem [46], the upper limit is $2n - 1$ (where n is the dimension of the state space) independent equations in this set. For bilinear state-space systems the upper bound (maximum number of independent equations) is $2^{2n} - 1$ [46] and for non-linear systems of homogenous polynomial form the corresponding upper bound is $(q^{2n} - 1)/(q - 1)$, where q is the degree of the polynomial [47], one selects (as a minimum) as many non-zero coefficients as there are basic parameters.

In summary, the procedure to apply the Taylor series approach for a general state-space model, involves the following main stages [42]:

(a) successive differentiation of $\mathbf{y}(t, \mathbf{p})$,

(b) evaluation of $y_i^{(k)}(0, \mathbf{p})$ for each $i = 1, \ldots, r$ by substitution of parameters already known from lower derivatives ($< k$) in addition to $y_i(0, \mathbf{p})$ and,

(c) a check on the parameters (from the parameter vector $\mathbf{p}$) that can be identified at each stage of the differentiation as well as a check on the independence of the algebraic equations in the successive derivatives.

In this thesis, the Taylor series approach is applied to the kinetic model in Chapter 4, the cell cycle model in Chapter 5 and the coupled drug kinetics/dynamics model in Chapter 6.

2.5 Parameter estimation

The main concern in the inverse problem is determining the system structure in addition to estimating the unknown parameters ($\mathbf{p}$) of the system. Estimating the values of the unknown parameters can also be considered to be part of verifying the structure of a state-space system. Once it has been ensured that the parameters in a model are identifiable (globally or locally) with respect to an experiment performed to collect data, one then must consider the problem of estimating the unknown parameters in the state-space model from these data. The effects of measurement errors, noise and imperfect data on parameter estimates have to be considered when performing parameter estimation. Additionally, it

has to be emphasised that there is no guarantee that globally (uniquely) identifiable parameters will be accurately estimated with high confidence from experimental data. Moreover, the precision (accuracy) of the estimated (fitted) values must be considered as a separate problem with respect to the experimental data.

There are many possible techniques for performing parameter estimation. Examples of such techniques include graphical methods (which are widely used) for estimating the components of exponential decay curves. However, graphical techniques work well only under three main conditions [23]; the underlying state-space model behaves in a linear fashion (or a linearised system), there are only a few exponential terms and the exponential decay constants, the eigenvalues (λ_i), are separated by factors of at least three. In graphical techniques, the accuracy of parameter estimates depends on the number of compartments in the model, if there are more than three exponentials (components), the possibility of obtaining accurate values is low. This method works well if there are no errors in the observations $\mathbf{y}(t, \mathbf{p})$, however, in real applications (i.e. parameter estimation from real data), there are always experimental errors. Hence, graphically derived parameter estimates have proven very good initial estimates for more powerful fitting algorithms including non-linear least squares techniques and direct estimation of basic parameters.

One of the most commonly used parameter estimation techniques for linear and non-linear state-space models is the direct estimation of basic parameters using software packages. This is performed by estimating the unknown model parameters directly by integrating the rate equations to obtain values of the observation function (output) using an initial estimate of the parameter values and then comparing the calculated values of the observation function (the output $\mathbf{y}(t, \mathbf{p})$) with the available data (measurements) at the sample points. Following that is moving in the parameter space in order to optimise the objective function. For each new set of parameter values, the rate equations have to be integrated, to generate the observation functions for this set of parameter values.

Parameter estimation in this thesis was performed using the commercial simulation software package FACSIMILE (MCPA Software, U.K.), which employs the VA05 routine from the Harwell Subroutine Library (a hybrid method combining features from the Newton-Raphson, Steepest descent and Levenberg-Marquardt techniques) [48]. This

software package is designed to numerically solve differential equations (rate equations), with specific focus on modelling the kinetics of physical and chemical models. A particular advantage of this computer-modelling tool is the robust (implicit) and reliable numerical integrator that is able to handle extremely stiff systems efficiently for both small and large dimensional problems. Stiffness is a property of the models presented in this thesis, that is, systems with widely varying rate constants, in other words, this problem appears when the system has some solutions that change very rapidly in comparison to other solutions or when solutions change very rapidly at some times compared with other times. The numerical integrator in FACSIMILE can solve all initial-valued ordinary differential equations within the given model simultaneously (where derivatives are approximated by differences over a finite time step) and uses the parameter-fitting option available to determine the unknown parameter values which best fit the simulated output to the measurements (experimental data). This computer-modelling tool uses a predictor-corrector method, in which the solution vector at the end of each time step is first predicted and then corrected using the hybrid method above, to satisfy the differential equations. The result of this optimisation method is the estimation of the unknown parameters of the model.

The parameter fitting is treated as an optimisation problem in which the set of unknown model parameters that minimises a weighted least-squares performance index is sought. That is, the performance of the optimisation process involves minimising the closeness of fit. This closeness (a measure of how well the calculated value fits the experimental data) is measured by the total weighted residual sum of squares (RSS), that is, the sum of the squared errors between a calculated value (model) and its corresponding data (observed) value divided by the standard error (error estimate for the data value) [49]. The parameter-fitting procedure involves a series of runs of the numerical integration process for the differential equations. The first run is conducted using the declared initial estimates for all unknown model parameters and then each unknown parameter is varied in turn (iteratively) and a statistical analysis is carried out to determine how well the calculated solution fits the observation data (experimental results). Simulations are carried out to vary the parameters until a set of unknown parameter values is estimated which gives a calculated solution that is qualitatively and quantitatively close to the experimental results. The integration process terminates when the highest value of the independent variable (in this case time t) has been

reached and when the fitting procedure has been completed, further simulation runs of the integration process are carried out to perform statistical analysis of the fitted parameters. In FACSIMILE the RSS is given by:

$$\mathrm{RSS} = \sum_{i=1}^{r} \sum_{j=1}^{N_i} \frac{(d_i(j) - y_i(t_j))^2}{\sigma_i^2} = \sum_{i=1}^{r} \sum_{j=1}^{N_i} G_{ij}^2 \tag{2.29}$$

where N_i is the number of actual observations used for the data curves $i = 1, \ldots, r$; $y_i(t_j)$ is the ith output of the model at the jth sampling time (t_j); and the corresponding experimental data point is denoted by $d_i(j)$ for each set of experimental data points (data curves). The estimated standard error associated with the data series $d_i = (d_i(1),\ldots, d_i(N_i))$ is denoted by σ_i [49]. The standard error of the output (σ_i) in FACSIMILE is equal to the product of e (estimate of the overall accuracy of the data) and R_i (range for the experimental data points) for each data curve, i.e., $\sigma_i = eR_i$. This approach is referred to as the weighted least-squares criterion since the random error at each sampling time (t_j) is multiplied by a constant weight ($1/\sigma_i^2$). This results in normalised residuals if the standard error of the output (σ_i) can be chosen to be the standard deviation (SD) for the random errors.

FACSIMILE estimates a range within which the RSS should fall (given that the number of data points is known), assuming that the model is accurate and the residual values are due to random errors in the experiment. If the final RSS is significantly smaller than the estimated (expected) range, then the fit is unpredictably good and therefore, the experimental errors might have been over-estimated. If the final RSS is remarkably greater than the estimated range, then the errors in the data might have been under-estimated or the model does not fit the experimental data very well (i.e., the fit is not close).

FACSIMILE performs other statistical tests to measure the quality of the fits including a correlation index for the residuals (CORI). This value for each data curve ($i = 1, \ldots, r$) is given by [49, 50]:

$$\mathrm{CORI}_i = \frac{\sum_{j=1}^{N_i} G_{ij}}{\sqrt{\sum_{j=1}^{N_i} G_{ij}^2}}. \tag{2.30}$$

If a good fit has been obtained, then the data values should be evenly distributed above and below the calculated values, i.e., the residuals are as likely to be positive as negative. If that is true, then $|\mathrm{CORI}_i| \approx 1$, otherwise, if the magnitude of CORI_i is far greater than 1.0 (i.e., $|\mathrm{CORI}_i| >> 1$) the residuals of the fit for data curve i are one-sided. Another statistical value that is calculated by FACSIMILE is the mean absolute correlation index ($\mathrm{CORI}_{\mathrm{mean}}$) which is given by:

$$\mathrm{CORI}_{\mathrm{mean}} = \frac{1}{r}\sum_{i=1}^{r}|\mathrm{CORI}_i|. \tag{2.31}$$

The method that FACSIMILE uses in estimating parameters involves the natural logarithms of the model parameters. During the fitting procedure, a statistical analysis is performed to detect parameters that are not well-determined (NWD) by the data. Once detected, NWD parameters are treated as unknowns in subsequent tests. The default standard deviation of the natural logarithm (SDLN) of each of the remaining fitted parameters, $\mathbf{P}^0$, is estimated from the variance-covariance matrix of $\mathbf{P} - \mathbf{P}^0$ [13]. The 5% and 95% confidence limits are estimated for each of the well-determined parameters, assuming a normal distribution for the natural logarithms of these parameters. The confidence limits are calculated using the SDLN value for each fitted parameter which are values such that the probability of the true parameter being smaller is 5% and 95% respectively [50]. If the 5% and 95% confidence limits for a particular parameter are reasonably close, this indicates that this parameter is well-determined. In addition, if the sampling errors are normally distributed, the weighted least-squares parameter estimate $\mathbf{P}^0$ is a maximum-likelihood estimate. Properties of maximum-likelihood estimators are discussed in [51], they are often unbiased, and if they are unbiased, maximum-likelihood estimators have minimum variance in comparison to all unbiased estimators [23]. However, if they are not unbiased, they become asymptotically unbiased as the size of the sample increases. FACSIMILE also calculates the components of the correlation matrix to show how the fitted parameters are related to each other.

Normal probability plots are used to test whether the weighted residuals (the weighted errors between the data points and the simulated output of the model), are normally distributed with standard deviation of 1 and zero mean [52]. To do this, the residuals are

listed in an ascending numerical order $R_1 < R_2 < \cdots < R_{N\times r}$ (the smallest value is numbered 1 and the largest is numbered $N\times r$), and

$$z_m = (m - 0.5)/(N\times r) \tag{2.32}$$

where (m = 1, …, $N\times r$), N is the number of sample points in each data curve and r is the number of data curves in each parameter-fitting process. Next, data tables for the normal distribution are used to obtain a Z_m value from the standard normal distribution corresponding to the cumulative probability (z_m) given in the previous step. Finally, the resulting graph from the Z_m values plotted against the ordered residuals R_m (horizontal direction) is the normal probability plot. If the residuals are normally distributed then the normal probability plot should be linear.

2.6 Summary

State-space modelling provides a framework under which drug kinetic and dynamic models can be described and analysed. This chapter provides the mathematical tools and techniques that will be used in the entire thesis for model development, model identification and parameter estimation. Structural identifiability analysis was performed in this thesis to estimate unknown model parameters with greater confidence. In addition, for each of the models described in this thesis, structural identifiability analysis was performed using MATHEMATICA, with the exception of the TPT hydrolysis model. In all models presented in this thesis, parameter estimation was performed using the software tool FACSIMILE.

3 Background biology and literature review

3.1 The anti-cancer agent topotecan

3.1.1 Introduction

Camptothecin (CPT), an ethanolic extract of the Chinese plant *Camptotheca acuminata* (Decaisne, Nyssaceae) [53, 54], which is also known as the Xi Shu tree, has revealed a noticeable anti-neoplastic activity, i.e., it inhibits or prevents the growth or development of malignant cells [55]. The anti-tumour action of CPT was established using experimental models during research conducted by the National Cancer Institute during a cytotoxic drug screening [56]. During clinical trials, CPT was tested in the form of water soluble carboxylate due to its limited solubility in water, in addition, undesirable side effects including myelosuppression (decrease in the activity of bone marrow) and hemorrhagic cystitis (infection of the bladder) have appeared during the trials [57]. However, the nuclear enzyme DNA topoisomerase I (Section 3.1.2) was discovered as the main CPT molecular target [58-62]. This finding re-initiated further development of CPT and its analogues as anti-cancer drugs [63-66]. Improved versions of the drug were developed using structure-activity data on CPT. Examples include topotecan (TPT), a water-soluble derivative [53, 54, 67], which has been clinically approved for use as a second-line therapy for small lung and advanced ovarian cancers [15, 68-70]. In addition, TPT has shown promising activity against haematological malignancies [67].

3.1.2 DNA topoisomerase I enzyme

A number of topological problems arise as a result of the intertwining of the two strands of the DNA helix [71]. Therefore, the cell must overcome such obstacles for it to be able to re-generate, re-combine and express its genetic information [72-74]. Topoisomerase I (or type I) is one of the two ubiquitous nuclear enzymes that modulate the topological

structure of DNA in eukaryotic cells; in effect, it interconverts different topological forms of DNA [71]. This enzyme has essential roles in cellular functions including chromatin organisation [75], DNA replication [76-79], recombination [80-82], mitosis [83] and transcription [78, 84-88]. In some types of tumours, the topoisomerase I enzyme is produced in excess when compared with corresponding non-malignant cells [70]. As mentioned in the previous section, topoisomerase I (Figure 3.1) is linked to the treatment of human cancers; type I is a primary cellular target for several types of anti-cancer drugs including CPT and its analogues.

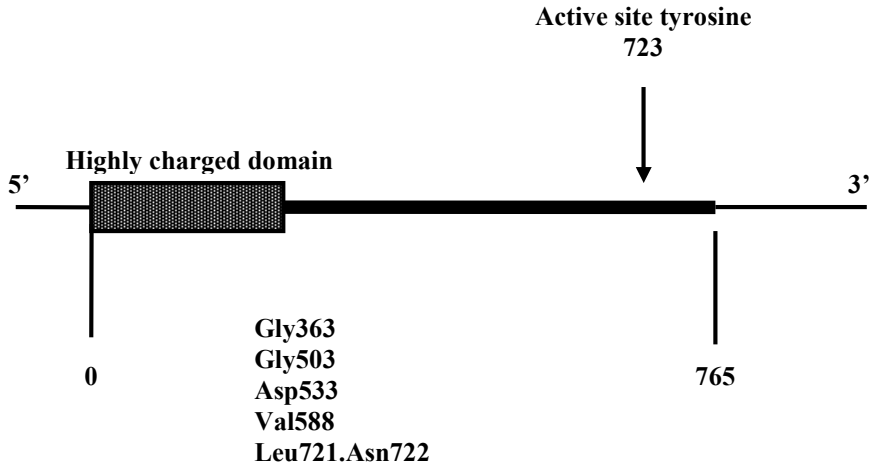


Figure 3.1: Human DNA topoisomerase I cDNA. The amino acid residues listed represent sites of mutation that can lead to CPT resistance (redrawn from Reference [53]).

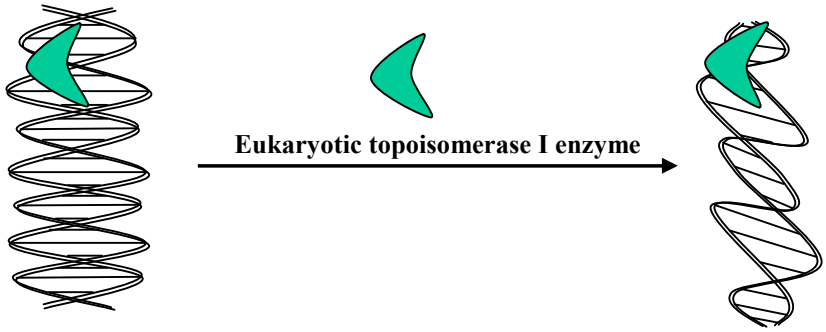


Figure 3.2: Super-coiled DNA and its relaxation by mammalian DNA topoisomerase I (redrawn from Reference [53]).

Excessive DNA supercoiling resulting from replication and transcription is removed by a relaxation reaction (Figure 3.2) of both positively charged and negatively charged supercoils [79, 89]. This function is catalysed by the topoisomerase I enzyme [53]. The

eukaryotic type I enzyme triggers a transient single-stranded nick in the nucleic acid backbone and passes the other DNA strand through the nick, and reseals the original scission [54, 72-74].

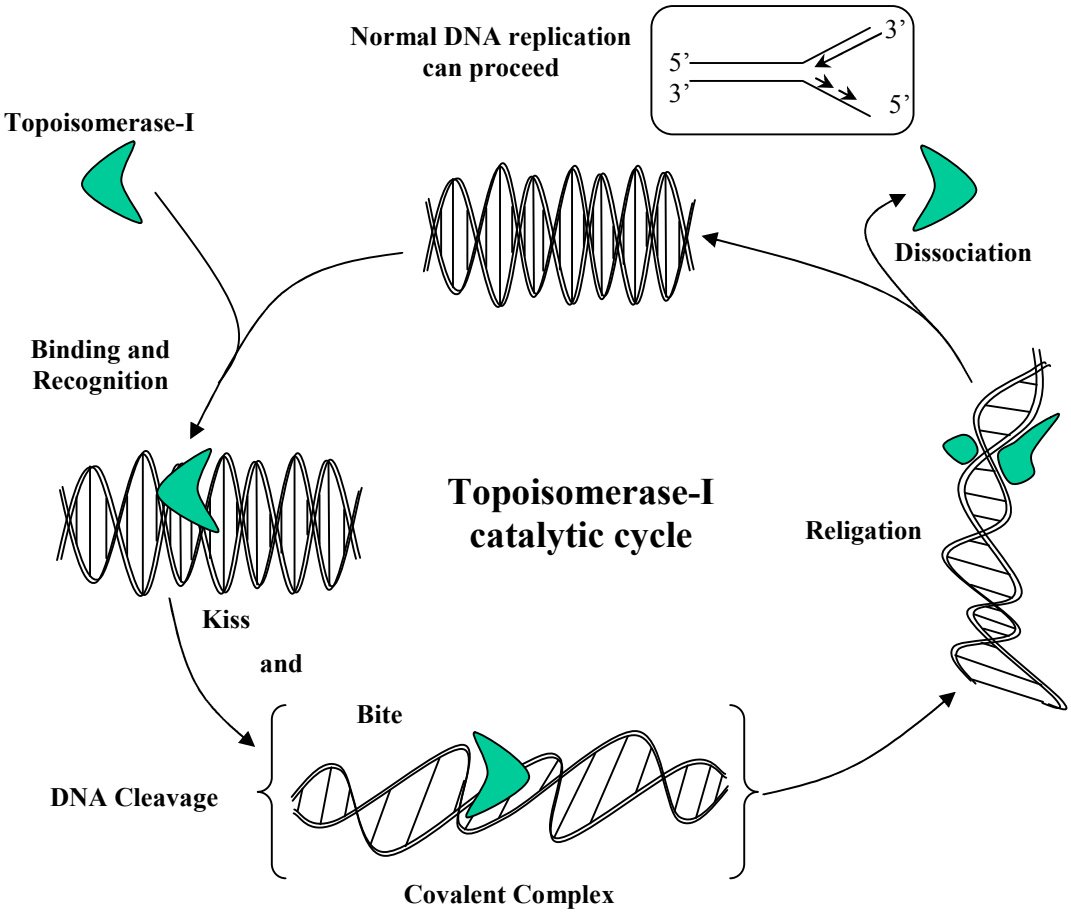


Figure 3.3: The topoisomerisation reaction during the cell cycle in the absence of TPT (redrawn from Reference [54]).

The schematic diagram in Figure 3.3 represents the reaction of the topoisomerase I enzyme with a DNA oligonucleotide. The catalytic cycle, which is also known as topoisomerisation, begins with the binding of the type I enzyme to the double-stranded DNA (dsDNA) resulting in a covalent 3'-phosphorotyrosyl adduct known as the cleavable complex [54]. As a result, the separated ends of the single-strand breaks are kept in proximity while the other DNA strand is passed through the break [90]. At this instant, the reaction is mainly controlled by the complementarity of the charge and surface of both species, i.e., the enzyme and the dsDNA. The shape adaptation of both partners results in the so-called *kissing* complex by which the tyrosine hydroxyl (OH) group at site Y723 [64]

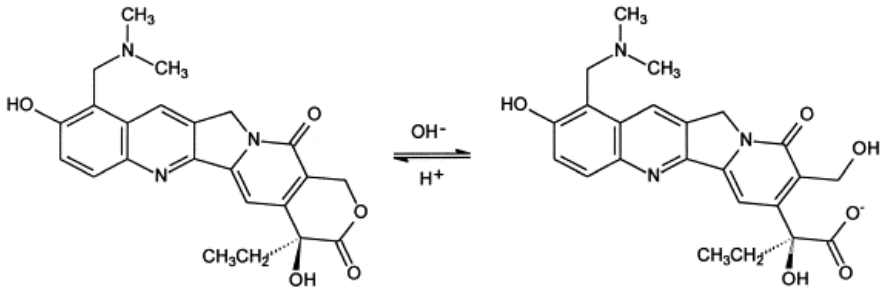
on the human topoisomerase I enzyme is brought into a favourable position for attacking the phosphate group of the phosphodiester linkage forming an enzyme-linked single-strand DNA break (covalent adduct) [91]. The *bite* of the enzyme topoisomerase I stimulates swivelling (the release of torsional stress in supercoiled DNA) in a process known as controlled rotation [92]. Next, the rejoining process begins by the 5' hydroxyl group attacking the tyrosol phosphate bond. As a result, the tyrosine residue is released from the phosphate and the single-strand break is re-ligated and the topoisomerase I enzyme dissociates from DNA, thus restoring the discontinuity of the DNA backbone and leaving the active site tyrosine ready for a later *kiss* and *bite* catalytic cycle. Single-strand breaks catalysed by topoisomerase I are barely detectable and are not lethal to cells since they can be well repaired.

3.1.3 TPT mechanism of action

CPT analogues including, but not limited to, TPT (Hycamtin®, GlaxoSmithKline) represent a new class of anti-cancer drugs that have been approved clinically for battling human cancers. These analogues have better water solubility and lower toxicities. In this project, the kinetics and dynamics of TPT will be discussed and studied. TPT has demonstrated clinical activity in human malignancies and it is one of the most widely used CPT analogues [69]. In addition, clinical trials are in progress to define its use in combination with other chemotherapeutic agents. Moreover, experiments are ongoing to better mark and describe its spectrum of clinical activity as well as to determine its optimal schedules of administration.

All CPT active analogues (including the active form of TPT) contain the chemically unstable lactone ring. Additionally, CPT analogues with more chemically stable rings are inactive (this includes the inactive form of TPT) [68]. Several studies have indicated that the administration of the closed ring lactone is essential for both topoisomerase I inhibition and maximal anti-tumour activity [93-96]. In aqueous media, the ring-closed (Figure 3.4(a)), parent lactone form (TPT_L) is chemically unstable at physiological pH and undergoes reversible hydrolysis to an open-ring hydroxy acid (hydroxyl carboxylate, Figure 3.4(b)) form (TPT_H), as shown in Figure 3.4, which is devoid of topoisomerase I

inhibitory activity [15, 97]. This conversion is pH-dependent, and under a high-acidic medium pH (≤4) the parent lactone structure predominates, whereas the hydroxy acid form is exclusively present in a high-basic medium pH (≥10) [98-100]. The pH has an effect on both, the lactone-to-carboxylate ratio at equilibrium as well as the conversion rate between TPT_L and TPT_H [98]. The study conducted by Yao *et al.* [101] has shown that a pH of 6.28 (at 25°C) resulted in an approximate 50:50 mixture of the two TPT species (TPT_L and TPT_H) and this agrees with the results obtained by Underberg *et al.* [98]. In addition, the rate of interconversion is increased with temperature rise without affecting the equilibrium itself *in vitro* and *in vivo*. In the latter, the equilibrium also depends on the binding process to albumin [102]. The reversible hydrolysis of TPT has been modelled by Evans *et al.* [15] (see Section 4.2) and forms the basic building block of the extended drug kinetics model presented in this thesis (see Chapter 4).



(a) **(b)**

Figure 3.4: Chemical structures of TPT (a) TPT_L and (b) TPT_H undergoing pH-dependent reversible hydrolysis (taken from Reference [103]).

The two forms of the anti-cancer drug TPT, i.e. TPT_L (lactone form) and TPT_H (hydroxy acid form), are two completely different chemical entities with significant differences in their physical properties. Therefore, they have different molecular interaction behaviours with target macromolecules including DNA, topoisomerase I and the topoisomerase I-DNA complex [101].

TPT_L is the pharmacologically active form of the drug that acts as a specific and reversible inhibitor of topoisomerase I [59, 96]. According to experiments, the lactone form (TPT_L) is much more cytotoxic than the carboxylate form (TPT_H) and the parent form also appears to be an important structural pre-requisite for passive diffusion of the anti-cancer agent into

malignant cells [104]. Therefore, it is essential for anti-tumour activity [93, 94, 96, 105]. The active form of the drug (TPT_L) targets the nuclear DNA-relaxing enzyme, topoisomerase I, and converts it into a cell poison by inhibiting the breakage-rejoining reaction of the enzyme. TPT_L covalently binds to DNA-topoisomerase I enzyme cleavable complex and stabilises it, specifically in the DNA re-joining (re-sealing) step of the reaction causing a delay (or blocking the re-joining step) [60, 106]. This results in accumulation of a reversible intermediate complex consisting of topoisomerase I + TPT_L + DNA as a ternary complex (cleavable complex, see Figure 3.5) [53]. The drug-induced single strand breaks are reversible upon drug removal. Such lesions are non-poisonous (non-toxic) as they can be reversed (repaired) [107]. TPT_L arrests the replication fork by stabilising the DNA-enzyme complex, and hence, causing double-strand breaks and terminating RNA-synthesis [108], this damage is responsible for the primary cytotoxic effects of TPT. The extent of DNA cleavage is reduced or increased depending on the point of the reaction cycle at which the anti-cancer drug TPT will act, whether it is before or after the *bite*.

It has been suggested that TPT_L only binds to the topoisomerase I-DNA complex and not to DNA alone nor to the topoisomerase I enzyme by itself as described by Chappell *et al.* [107]. However, the data published by Yang *et al.* [109] show that the active form of TPT is stabilised through direct interactions with dsDNA. In addition, the presence of dsDNA was found to promote the conversion of TPT_H to TPT_L (within the cell nucleus) in the absence of topoisomerase I enzyme. This suggests that TPT, upon reaching chromosomal DNA, may interact directly with DNA prior to the action of the topoisomerase I enzyme even if the drug binding site on DNA is not necessarily at the site of the enzyme's (topoisomerase I) action [109]. In the study published by Yao *et al.* [101] it was observed, for the first time, that only TPT_L binds directly and non-covalently to both dsDNA and single-stranded DNA (ssDNA) structures in the absence of the topoisomerase I enzyme. Binding to dsDNA and ssDNA is selective. Additionally, it has been demonstrated that TPT_L exhibits specificity for dT-containing sequences in dsDNA and exhibits a characteristic binding pattern to ssDNA polymers [101].

Some malignant tissues contain higher topoisomerase I levels of expression than their normal counterparts [110, 111]. In addition, different tumour types have different

expression levels of the enzyme, for example, higher expression of topoisomerase I has been detected in colon and cervical cancers than in lung and breast tumours [112]. The level of topoisomerase I is relatively constant in between the cell cycle phases [108]. Moreover, the level of type I enzyme is regulated at transcriptional, translational and post-translational levels [113, 114]. However, this requirement for replication fork movement to produce cytotoxicity implies that TPT is a cell cycle specific drug for killing cells in synthesis (S) phase and causing extensive chromosomal disruptions and chromatid exchanges [108, 115, 116]. This property has significant implications for the clinical use of this anti-cancer agent, because optimal therapeutic efficacy of S phase-specific cytotoxic drugs (including TPT) requires protracted exposure of the tumour to drug concentrations exceeding a minimum threshold [67]. In fact, studies involving *in vitro* and *in vivo* experiments suggest that for efficacy (anti-tumour activity or cytotoxicity), prolonged administration of the anti-cancer drug TPT is of greater significance than short-term exposure to high concentration [117].

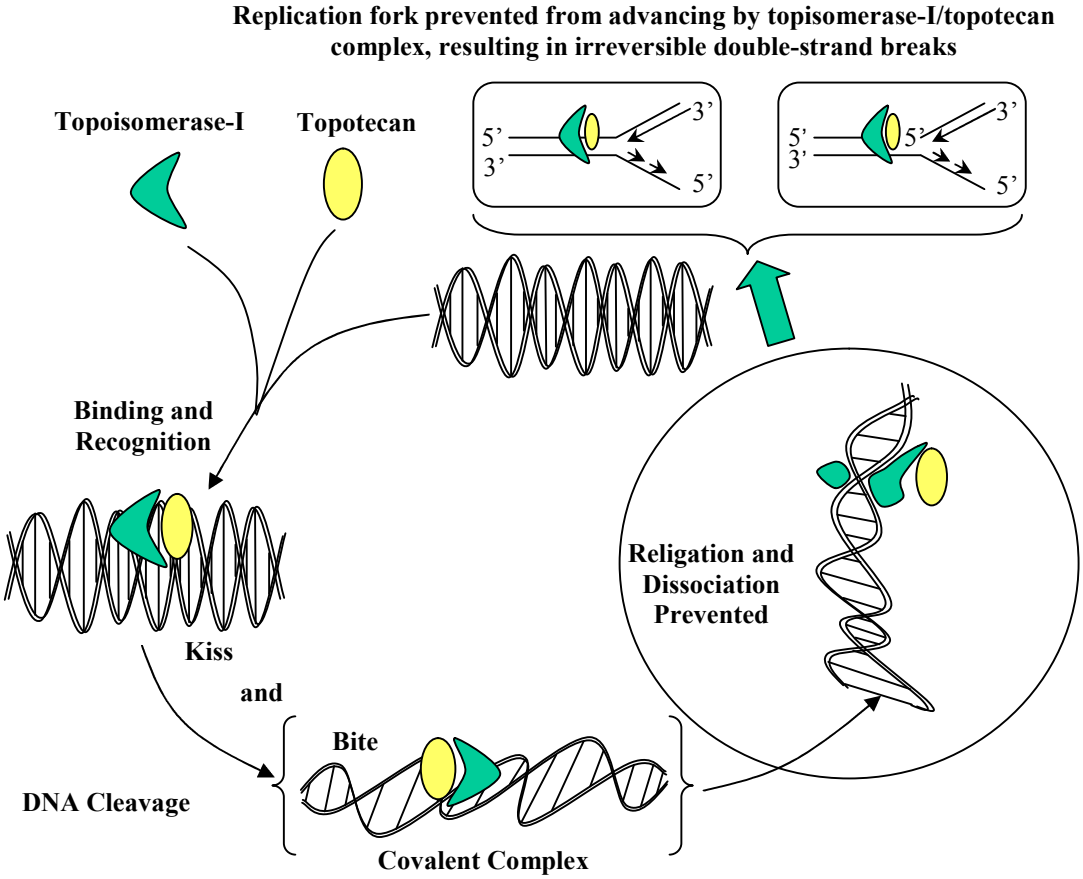


Figure 3.5: The topoisomerisation reaction during the cell cycle in the presence of TPT (redrawn from Reference [54]).

3.2 Drug resistance mechanisms

3.2.1 The BCRP/ABCG2 drug efflux pump

One of the major forms of cancer treatment is chemotherapy. In general, cancers are either resistant to chemotherapy or obtain resistance during treatment, thereby leading to ineffective chemotherapy [118]. Several resistance mechanisms to topoisomerase I-targeted anti-cancer agents have been characterised *in vitro*, however, their significance in the clinical setting has not been fully identified until now. These mechanisms involve pre-target events such as drug accumulation, metabolism, intracellular drug distribution, or drug-target interactions [67]. Post-target events, including DNA synthesis or repair, cell cycle progression as well as cell death regulation, have also been demonstrated to play a significant role in the sensitivity to these drugs [67].

Numerous mechanisms have been described in cancer cells able to reduce accumulation of cytotoxic drugs by up-regulating cell membrane located adenosine triphosphate (ATP)-dependent efflux pumps [119]. Examples of drugs to which resistance has been observed or for which resistance is acquired during treatment include topoisomerase inhibitors (e.g., TPT), anthracyclines (e.g., doxorubicin), the *Vinca* alkaloids (e.g., vincristine) and the taxanes (e.g., paclitaxel) [118]. The model in this thesis assumes that a resistance to TPT is a result of the over-expression of efflux transport proteins in the plasma membrane of cancer cells. Examples of such mechanisms include the two transmembrane xenobiotic transporter proteins, P-glycoprotein (P-gp) which has been reported in hamster ovarian cells [120, 121], and the multidrug resistance protein 1 (MRP1) [122]. More recently, other human ATP-binding cassette (ABC) transporters that are involved in the resistance of anti-cancer drugs have been discovered. The most important of these, a novel protein, is the breast cancer resistance protein (BCRP) [123] also known as the placenta-specific ABC protein (ABCP) [124] or mitoxantrone-resistance protein (MXR) [125].

Following immunoblotting analysis with BCRP-specific antibodies, it has been suggested that BCRP is a 72 kDa membrane protein [126, 127]. It has been shown from various

studies on drug efflux mechanisms that BCRP is mostly localised at the plasma membrane of the drug-resistant cells over-expressing the transporter instead of at internal vesicular membranes [123, 128, 129]. Therefore, BCRP is presumably involved in active transport from the cell rather than in transport into internal vesicles. Following experimental studies by Maliepaard *et al.* [130] and Ishii *et al.* [131] it has been observed that TPT accumulation in resistant cell lines was reduced by enhanced drug efflux. The actual efflux pumping mechanism is explained in detail in Chapter 4. Transfection studies conducted in various laboratories have verified that enforced expression of BCRP complementary DNA (cDNA) in different types of cells caused resistance to multiple anti-cancer drugs and reduced drug accumulation in the cell [128, 129, 132]. Such experiments provide strong evidence that BCRP is a main cause of drug resistance in tissue culture models for certain types of chemotherapeutic drugs including TPT. This protein is significantly expressed in organs central for absorption (the small intestine), distribution (the placenta and blood-brain barrier), and elimination (the liver and small intestine) [118]. Further investigations have come to the conclusion that BCRP plays a significant role in drug disposition.

The ABC transporter is the largest protein superfamily identified to date [133]. ABC transporters are widely spread in all organisms (including mammals and bacteria) and are responsible for transporting a wide range of compounds across the plasma membranes against concentration gradient (active pumping mechanism) with ATP hydrolysis as the source of energy for the process of substrate translocation. ABC transporters are involved in multiple physiological processes. Examples of this include, transporting drug (xenobiotics) or drug conjugates in addition to the excretion of endogenous metabolites or physiological substrates [134]. A number of human genetic diseases have been found to be associated with defects in ABC transporter genes such as CFTR in cystic fibrosis [135], MRP2 in Dubin-Johnson syndrome [136] and ABCR in Stargardt disease [137]. Typically, the structure (Figure 3.6) of the majority of mammalian ABC transporters contains two types of structural domains (homologous halves) each containing two parts as follows: the hydrophilic intracellular nucleotide binding domain (NBD) and the hydrophobic membrane spanning domain (MSD) which is putatively arranged into six α-helices [118, 138]. P-gp is arranged in two tandem repeated halves of two domains: one hydrophobic MSD followed by one hydrophilic NBD [120, 134]. The two repeated halves are connected

by a polypeptide linker sequence. The great majority of ABC transporters have such a 4-domain structural organisation.

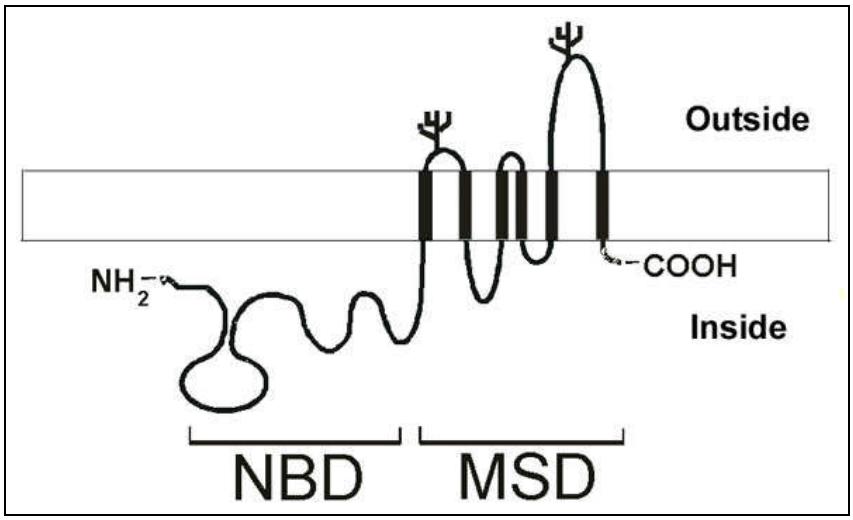


Figure 3.6: A membrane topology model of BCRP. BCRP contains one NBD followed by one MSD with six predicted transmembrane α-helices. Two or three putative N-glycosylation sites (N418, N557, or N596) are predicted to be in the extracellular loops as indicated (taken from Reference [118]).

BCRP belongs to the sub-family G of the large ABC transporter superfamily. BCRP, also known as ABCG2, is the second member of the sub-family G of the human ABC transporter. This transporter has been described in drug-resistant ovary, breast, colon, gastric cancer in addition to fibrosarcoma cell lines [123-125, 130, 139]. Moreover, BCRP has only one ATP-binding cassette and six putative transmembrane domains, suggesting that BCRP is a half-transporter, which may work as a homodimer or heterodimer [140, 141], unlike P-gp and MRP1 which are arranged in two repeated halves. Another unique feature is the configuration of the (BCRP/ABCG2) protein in which the NBD is followed by the MSD as illustrated in Figure 3.6, whereas P-gp and MRP1 have an opposite domain arrangement, in that the MSD precedes the NBD. Therefore, this unique domain organisation implies that the transport mechanism of ABCG proteins (including BCRP/ABCG2) may be different from those of other ABC transporters.

CPT derivatives are the second most important class of anti-cancer agents that are transported by the (BCRP/ABCG2) transporter [118]. A wide range of BCRP-over-expressing cell lines reveal resistance to CPT derivatives including TPT [131, 142]. The

study carried out by Rabindran *et al.* [143] demonstrates that MCF-7 cells transfected with BCRP cDNA display significant resistance to TPT in comparison to the vector control cells, this is additional evidence that CPT derivatives (including TPT) are BCRP substrates. In a study published by Rajendra *et al.* [144], it has been reported that the important factors for substrate recognition (i.e. effective efflux pumping) are the hydrophilic groups in the chemical structure of CPT. It has been shown by Yoshikawa *et al.* [145] that BCRP prefers to transport CPT derivatives with high polarity over the low polarity CPT analogues [118]. Accordingly, polarity is also essential for recognition of the CPT analogues by BCRP. All such information is essential in designing clinically useful CPT analogues that are not transported by BCRP/ABCG2. Based on the advice of the project collaborators at Cardiff University and according to these observations, it is assumed that in the drug kinetics model (see Chapter 4) the BCRP/ABCG2 transporter carries TPT_H (high polarity) from the cytoplasm through the cell membrane to the extracellular region irreversibly.

3.2.2 The enzyme ALDH

Aldehyde dehydrogenases (ALDHs) [aldehyde: $NAD(P)^+$ oxidoreductase] are an enzyme group that catalyse the conversion of aldehydes to the corresponding acids irreversibly by means of an $NAD(P)^+$-dependent reaction [146]. ALDHs are distinguished based on properties including physicochemical characteristics, enzymological, subcellular localisation and tissue distribution [147]. ALDH enzymes demonstrate high activity in oxidising aldophosphamide and detoxifying anti-cancer agents. It has been proven that this drug resistance was associated with the transcriptional activation of ALDH expression in cancer cells [148].

ALDH isozymes[6] (or isoenzymes) including ALDH1A1, are important for multiple biological activities including drug resistance [149, 150]. It has been demonstrated by the project collaborators at Cardiff University that the uptake of TPT in cells exposed to an ALDH inhibitor (disulfiram) is enhanced (unpublished; data not shown). Moreover, kinetic

[6] Isozymes: are enzymes that vary in the amino acid sequence but catalyse the same chemical reaction.

analysis of TPT uptake also suggests that the change in charge upon ring hydrolysis would favour exclusion via the ABCG2 pump, and this is the overriding process to remove TPT from the cellular environment [151]. Therefore based on these studies, the active drug (TPT_L) in the cytoplasm is assumed to bind irreversibly to the enzyme ALDH (in the extended drug kinetics model) where it is converted to the form favoured by the efflux pump (TPT_H).

3.2.3 Fluorescence spectral properties of TPT

TPT is a naturally fluorescent agent which can be excited using ultra-violet wavelengths (350-430 nm) [100] and detected at a visible wavelength with peak emission at 540 nm. This property of TPT has been exploited to study the interaction of this anti-cancer drug with biological macromolecules and most importantly to evaluate drug resistance in differentially derived cell lines using confocal microscopy [152]. This approach is commonly used to compute the amount of drug in human plasma and cellular extracts [15] using HPLC [103]. The binding characteristics of drug to DNA have been determined using the high two-photon absorption property (demonstrated by TPT) during spectroscopy studies [100].

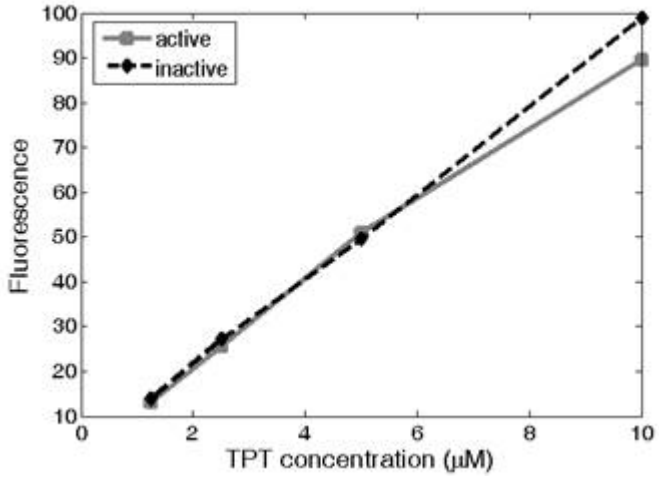


Figure 3.7: The fluorescence intensity of TPT in live human breast cancer cells (MCF-7 cell line) plotted against the concentrations of TPT_L and TPT_H (taken from Reference [48]).

The anti-cancer drug TPT is also infrared two-photon excitable; the basic principle of two-photon excitation of fluorescence is that two photons of longer wavelength light are absorbed (simultaneously) by a fluorochrome (a chemical entity that exhibits fluorescence) that would generally be excited by a single photon with a shorter wavelength. Accordingly, the excitation of fluorochrome is limited to the focus point by the non-linear optical absorption property of two-photon excitation. Advantages of such a method (i.e., non-linear excitation) include the following: images obtained have relatively high contrast and are free of out-of-focus light. In addition, this technique is beneficial to live cell imaging since the light is restricted to the focal plane. Accordingly, this decreases the photo-damage to the cells as well as photo-bleaching of the indicator. This method is however not capable of differentiating between both forms of the drug, i.e. TPT_L (active form) and TPT_H (inactive form) as demonstrated in Figure 3.7. In the experiments conducted at Cardiff, delivery and localisation of the drug in single MCF-7 cells (live human breast cancer cells) are examined using TPLSM (see Subsection 4.5.1).

3.3 The cell cycle

3.3.1 Introduction

Recent studies have examined tumours and tumour cell lines for the presence of mutations in the genes encoding cell cycle-related regulatory proteins [153]. The results of these studies demonstrate that mutations in the cell cycle genes are the most common genetic alterations in cancer cells [154]. Therefore, the cell cycle has been studied extensively in cancer research (oncology). This research has led to techniques useful for determining the effects (pharmacodynamics) of anti-cancer drugs on the cell cycle progression. Hence, the cell cycle response (by mathematically modelling the dynamics of anti-cancer drugs) should be considered in the design of chemotherapeutics. In this thesis, a model describing the response of the growth of single human cells in the absence (see Chapter 5) and presence (see Chapter 6) of the anti-cancer agent TPT is presented.

The cell cycle is probably the most essential process that occurs in eukaryotic cells [155]. It is a ubiquitous, complicated process engaged in proliferation and growth of cells. The cell cycle is also involved in organism development, regulation of DNA damage repair, tissue hyperplasia (hypergenesis) as response to chronic inflammations and injuries [156]. In addition, this process is associated in life-threatening diseases such as cancer.

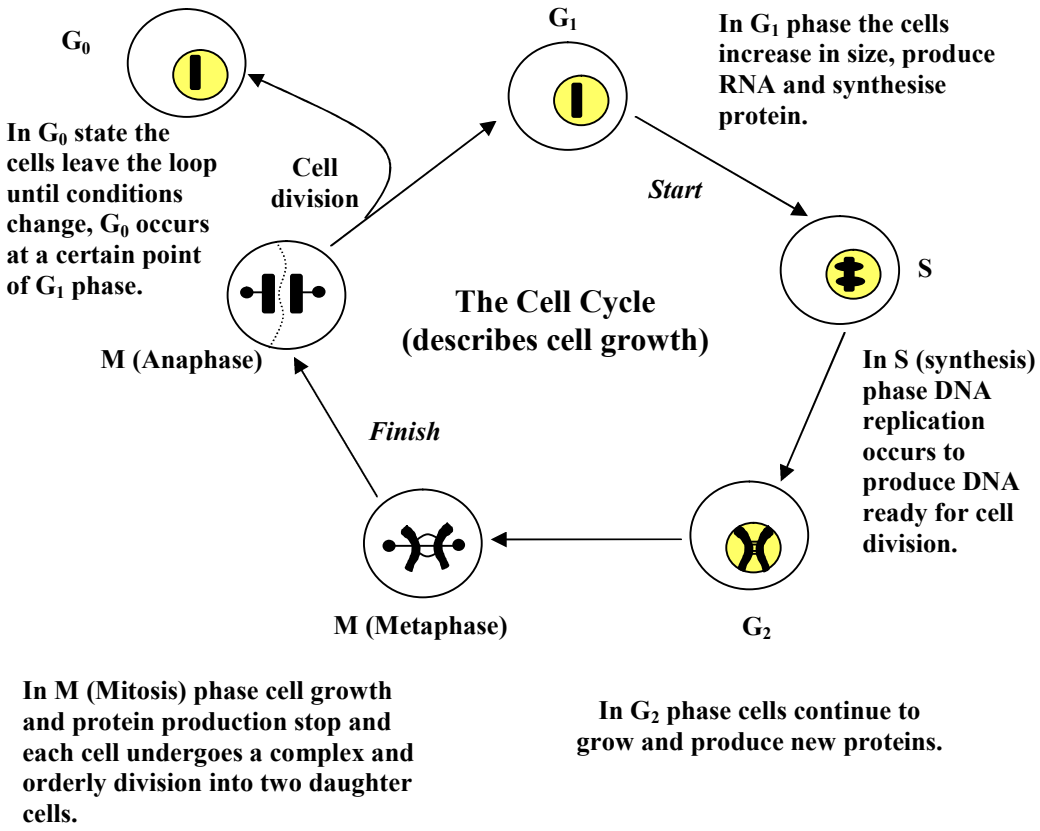
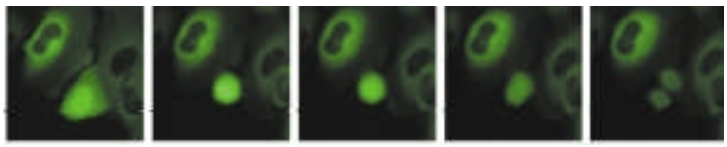


Figure 3.8: Schematic of the cell cycle showing the main phases and the *Start* and *Finish* transitions (redrawn from Reference [107]).

The cell cycle (cell division cycle) is the series of events by which a growing cell replicates all its essential constituents (components) and divides into two daughter cells. Each of the two distinct cells contains adequate machinery and information to repeat the process [157]. The most important components are the chromosomes of the cell which contain linear DNA molecules associated with proteins. Each DNA molecule must be precisely duplicated and the two copies accurately partitioned to the two daughter cells at division [158]. In mammalian somatic cells, these processes occur in temporally distinct phases as shown in Figure 3.8. The cell cycle comprises four stages (phases) referred to as G_1 (Gap 1), S (Synthesis), G_2 (Gap 2) and M (Mitotic) phases. The two milestones in this process

are DNA synthesis and mitosis (which also make up the chromosome cycle of the cell), and the gaps (G_1 and G_2) of the cycle occur between these two landmarks.

In G_1 stage, the chromosomes are unreplicated and the cell is not committed to division, moreover, cells in this phase start preparing for DNA synthesis. During transition to S phase, the cell commits to replication if the integrity of DNA is not compromised; that is, the cell undergoing the cycle confirms that the internal and external conditions are suitable for a new round of DNA synthesis. In S phase, the cell starts synthesising DNA and therefore, it includes aneuploid (between 2N and 4N) DNA content. The two identical DNA molecules are referred to as sister chromatids. This is followed by the G_2 stage during which the cell prepares for the mitotic process (M phase) where the chromosomal material is prepared by the surveillance of its integrity before committing to mitosis. For example, in the G_2 phase, the cell checks that the DNA synthesis and repair are complete before committing to the M phase. Additionally, persistence of double strand breaks that might be caused by an anti-cancer drug (see Section 3.1) are sensed at the cell cycle checkpoints.



(a) **(b)** **(c)** **(d)** **(e)**

Figure 3.9: The four stages of the M phase. (a) Prophase, (b) metaphase, (c) anaphase (d) telophase followed by (e) cytokinesis (taken from Reference [159]).

The mitotic (M) phase in the cell cycle includes mitosis and cytokinesis (the true division of cells into two identical daughter cells). The M phase can be subdivided into four stages: prophase, metaphase, anaphase and telophase, which are followed by cytokinesis [159, 160]. The mitotic process [161] as shown in Figure 3.9, begins with prophase; during this stage, the nucleoli gradually disappear and chromatin (the chromosomal material) begins to coil into dense, recognisable chromosomes. Each replicated chromosome consists of two chromatids, which both carry identical genetic information. Also during this phase (in some texts it is referred to as prometaphase), the nuclear membrane (envelope) fades away and the nucleus becomes unrecognisable. Moreover, the spindle fibres start forming;

spindle fibres are fibrelike processes that extend from the poles toward the equatorial plate of the cell. Some fibres elongate from centrosomes and fasten to kinetochores (complex protein structures located at the centromere region in the chromosomes that specify the attachment between the two sister chromatids and the spindle fibre microtubules [162]). Other spindle fibres elongate and overlap each other at the centre of the cell instead of attaching to the chromosomes.

In metaphase, the next mitotic phase, chromosomes are aligned in one plane along the polar axis as a result of the tension applied by the spindle fibres. Anaphase follows metaphase, and is characterised by the shortening of the mitotic spindle fibres and separation of the kinetochores so that the chromosomes (formerly chromatids) are pulled apart and begin moving to opposite poles of the cell along the direction established by the spindle fibres. The attachment of kinetochores to the spindle fibres is essential for the movement to take place in the anaphase stage. If the kinetochore is lost, the sister chromatids will not be able to orient on the spindle fibres, and as a result, the separation fails to occur. The next stage is the telophase; during this phase, the daughter chromosomes (formerly sister chromatids) arrive at the poles and the mitotic spindle fibres that have pulled them apart fade away. The final phase is cytokinesis in which spindle fibres not attached to chromosomes disappear and the cell also undergoes the furrowing process in which the cell squeezes-in on all sides until two daughter cells are formed. In addition, a nuclear membrane is reformed, the chromosomes uncoil and disappear (under the light microscope) as discrete structures and the cell returns to the interphase (G_1) state.

Cells in the G_0 (resting or quiescent) state are not actively cycling. The resting state is used to indicate cells not in the cycle except with the potential for division; examples include hepatocytes (cells in the liver involved in protein synthesis, protein storage and other vital activities). G_0 state occurs at a certain point in the G_1 phase, in this state, most of the cell functions are halted (including proliferation).

The growth cycle runs parallel to the cell cycle, whereby the components of the cell (including proteins, (ribonucleic acid) RNA, carbohydrates and phospholipid bilayers) are also replicated and partitioned (after cytokinesis) more-or-less evenly between the two

daughter cells. These two cycles (i.e., the cell cycle and growth cycle) turn at the same rate during normal cell proliferation, so that each round of DNA synthesis and mitosis is balanced by duplicating all other macromolecules in the cell. Therefore, the nucleocytoplasmic ratio (ratio of nuclear size to the cell size) is kept within advantageous limits. However, there are exceptions to this rule, for example, fertilised eggs divide rapidly without growing and oocytes grow largely without dividing [163]. Nonetheless, balanced growth and division determine the long-term viability of a cell line.

The major processes of the cell cycle are tightly regulated. In most eukaryotic cells this coordination is carried out at three checkpoints [158]. A checkpoint (regulatory mechanism) is a stage of the cell cycle process where the progress is stopped until the appropriate relay signals (surveillance mechanism) confirm suitable conditions for progress throughout the cycle. For example in G_1 phase, cells explore their internal state (nucleocytoplasmic ratio) and external environment (for nutritional conditions and growth factors [158]) and this is known as the G_1/S checkpoint. That is, for the *Start* transition (see Figure 3.8) to happen, cells must grow to a critical size before they can commit to DNA synthesis and division to achieve balanced growth and division. However, this requirement can be compromised by mutations and as a result, cells may become morbidly large or small [164]. In G_2 phase, the cell checks DNA synthesis and repair and that the size is large enough to commit to the M phase; this surveillance mechanism is known as the G_2/M checkpoint. In M phase (particularly at metaphase), cells ensure that all chromosomes are properly aligned on the metaphase plate and that there are no problems with DNA replication before issuing the command to segregate sister chromatids to opposite poles of the mitotic spindle. This is a crucial regulatory constraint, if there have been any problems, the *Finish* (see Figure 3.8) transition is held off at the third checkpoint (mitotic checkpoint). If anaphase commences without fulfilling the conditions of the mitotic checkpoint, a fatal mistake might happen and the daughter nuclei would not receive a full complement of chromosomes [165].

For modelling purposes, according to Tyson and Novak [157] the cell cycle can be subdivided into two alternative states, G_1 and S-G_2-M separated by *Start* and *Finish* as in Figure 3.8. *Start* (the transition from G_1 phase to S phase) is defined as when the cell begins replication, the decision is irreversible, i.e. once DNA synthesis begins, it goes to

completion [157]. The second irreversible transition (i.e., *Finish*) is triggered when the DNA replication is complete. The anti-cancer agent TPT is believed to be S phase specific, this implies that many cells may resist the effects of the drug by not attempting to replicate DNA (i.e., failing to enter the S phase) during a finite exposure to TPT [166]. Blocking entry to S phase results in suppressing cellular growth and expansion (i.e., cytostasis) and under some conditions can cause apoptosis (i.e. cell death).

3.3.2 Cell cycle regulation

Cyclin B1 is a protein involved in the regulation of the cell cycle. In this thesis, a mathematical model of Cyclin B1 dynamics is extended from a previously published model by Pomerening *et al.* [14] describing the growth of human osteosarcoma cells (U2-OS cell) in the absence (see Chapter 5) and presence (see Chapter 6) of the anti-cancer agent TPT. The model accounts for the inhibition dynamics of the cyclin-dependent kinase inhibitor (CKI) $p21^{CIP1/WAF1}$ on Cyclin B1.

Numerous regulatory proteins are involved in the cell cycle; in fact, they direct the cell through specific series of events as well as terminating in mitosis and producing two daughter cells. The cell cycle progression is regulated (controlled) by cyclin-dependent kinases (CDKs) and their activating partner proteins, cyclins [167]. CDKs are critical for the cell cycle progression because their inactivation prevents mitosis [168-170]. The activation of CDKs necessitates binding to a cyclin. Moreover, CDKs are targeted to the nucleus when bound to cyclins, since CDKs lack nuclear localisation signals that are contained in cyclins [171]. Some CDKs attach more than one cyclin, hence this provides some substrate specificity. The expression of cyclins during the cell cycle process is cyclic, that is why they were so named. Therefore, CDKs are activated at specific times during the cell cycle as a result of this cyclic expression.

There are several CDKs in eukaryotic cells; CDK1, also known as CDC2 (this designation used to name cell division cycle genes and proteins), is the prototype CDK. Cyclin B1 is one of the most extensively characterised cell cycle regulators in human and other

eukaryotic cells [159]. In the G_1 phase, CDK1 activity is relatively low because the relevant partner cyclin protein (Cyclin B1) is missing. Cyclin B1 is missing during G_1 because Cyclin B1 mRNA production is inhibited and Cyclin B1 protein is degraded. At *Start* (late G_1) Cyclin B1 synthesis is promoted and Cyclin B1 degradation is inhibited causing a remarkable rise in the activity of CDK1. CDK1/Cyclin B1 complex activity remains high throughout S, G_2 and M phases of the cell cycle. High activity of CDK1/Cyclin B1 dimer is necessary for DNA replication, chromosome condensation and spindle assembly.

At *Finish* (late G_2) the anaphase-promoting complex (APC) is activated. APC consists of a group of proteins that targets proteins such as Cyclin B1 for degradation by attaching a destruction label. The target proteins are degraded by proteolysis (digestion of proteins by cellular enzymes). The core complex of APC is made up of about a dozen polypeptides in addition to two auxiliary proteins, namely CDC20 and CDH1. The main roles of CDC20 and CDH1 are to recognise specific target proteins including the CDK1/Cyclin B1 dimer and present them to the core complex to be labelled for destruction [172, 173].

There are other complexities in the CDK network and these proteins (i.e., CDKs) are in turn regulated by numerous proteins such as the CKIs including p53 and $p21^{CIP1/WAF1}$. The cell cycle is often dysregulated in tumour cells (neoplasia) as a result of alterations in tumour suppressor genes or oncogenes that have a direct affect on cell cycle regulation such as p53 in addition to oncogenes that indirectly impact the cell cycle process [156]. p53 is a tumour suppressor gene and mutations of this gene have been identified in many types of malignancies including osteosarcomas [174-177]. This tumour suppressor gene controls multiple cell cycle checkpoints that regulate mammalian response to DNA damage. Following DNA damage, the tumour suppressor gene is stabilised and initiates the transcription of the $p21^{CIP1/WAF1}$ protein resulting in cell cycle arrest (or apoptosis) in the G_1 stage of the cell cycle. That is, p53 halts G_1/S (*Start*) transition in cells exposed to DNA-damaging agent by causing accumulation of $p21^{CIP1/WAF1}$ [178, 179], a protein that binds to CDK/cyclin complexes (including CDK/Cyclin B1 complex) [180] to inhibit their activity for initiating DNA synthesis [181]. p53 is also involved in regulating a mitotic spindle checkpoint that prevents DNA synthesis prior to chromosome segregation [182]. The CKI $p21^{CIP1/WAF1}$ can also bind to proliferating cell nuclear antigen and inhibit its

function in replication, this process also results in G_1 arrest. Moreover, $p21^{CIP1/WAF1}$ can prevent G_2/M transition (under the transcriptional control of p53) in human cell lines [183-185] by decreasing the intracellular levels of Cyclin B1 protein (the regulatory subunit of CDK1) which is required for mitotic initiation [186]. As a result, p53 (by expressing $p21^{CIP1/WAF1}$) arrests the cell cycle process in G_2 phase [187]. The ability of the tumour suppressor gene p53 to control mitotic initiation (by inducing the expression of $p21^{CIP1/WAF1}$) by regulating the intracellular levels of Cyclin B1 protein suggests that Cyclin B1-dependent G_2 checkpoint is involved in preventing growth of malignancies, i.e. preventing cells with damaged DNA from entering the M phase of the cell cycle [186].

Since the progression of the cell cycle is tightly regulated by several proteins, including Cyclin B1, for that reason, Cyclin B1 protein tracking provides information on the cell cycle progression and the dynamics of the cell regulators. In this work, the monitoring of the cell cycle progression was achieved by the project collaborators (Wales School of Medicine, Cardiff University) using the green fluorescent protein (GFP)-probe (see Chapter 5) which has expression, localisation and destruction description that shadow Cyclin B1 dynamics in living cells [159].

The cell cycle regulation is also determined by phosphorylation state [156]. Phosphorylation on threonine and tyrosine residues is one of the methods in which the CDKs are regulated. CDK1 has specific phosphorylation sites that are both stimulatory and inhibitory depending on the location of phosphorylation on CDK1 residues. For example, CDK1 phosphorylation on threonine 161 (T161) by cdk-activating kinase (CAK) is essential for CDK1 kinase activity, and this is a stimulatory function [188]. In addition, threonine 14 (T14) and tyrosine 15 (Y15) are located in the active site of the kinase and their phosphorylation prevents kinase activity. This phosphorylation is performed by Wee1, Mik1 and Myt1 protein kinases [169, 189-194]. The Wee1 kinase phosphorylates and inactivates CDK1/Cyclin B1 complex upon its entry into the nucleus, in order to protect the nucleus from premature exposure to an active mitotic kinase [190]. When the cell becomes ready to divide, CDC25 dephosphorylates CDK1 at T14 and Y15, and therefore activating CDK1 kinase [195-198]. Wee1 kinase and CDC25 phosphatase are in turn regulated by phosphorylation, therefore forming an activation cascade [199-201]. Moreover, CDK1 phosphorylates both Wee 1 and CDC25, thereby forming positive- and

negative-feedback loops as shown in Figure 3.10. These feedback components are incorporated in the extended cell cycle model presented in Chapter 5.

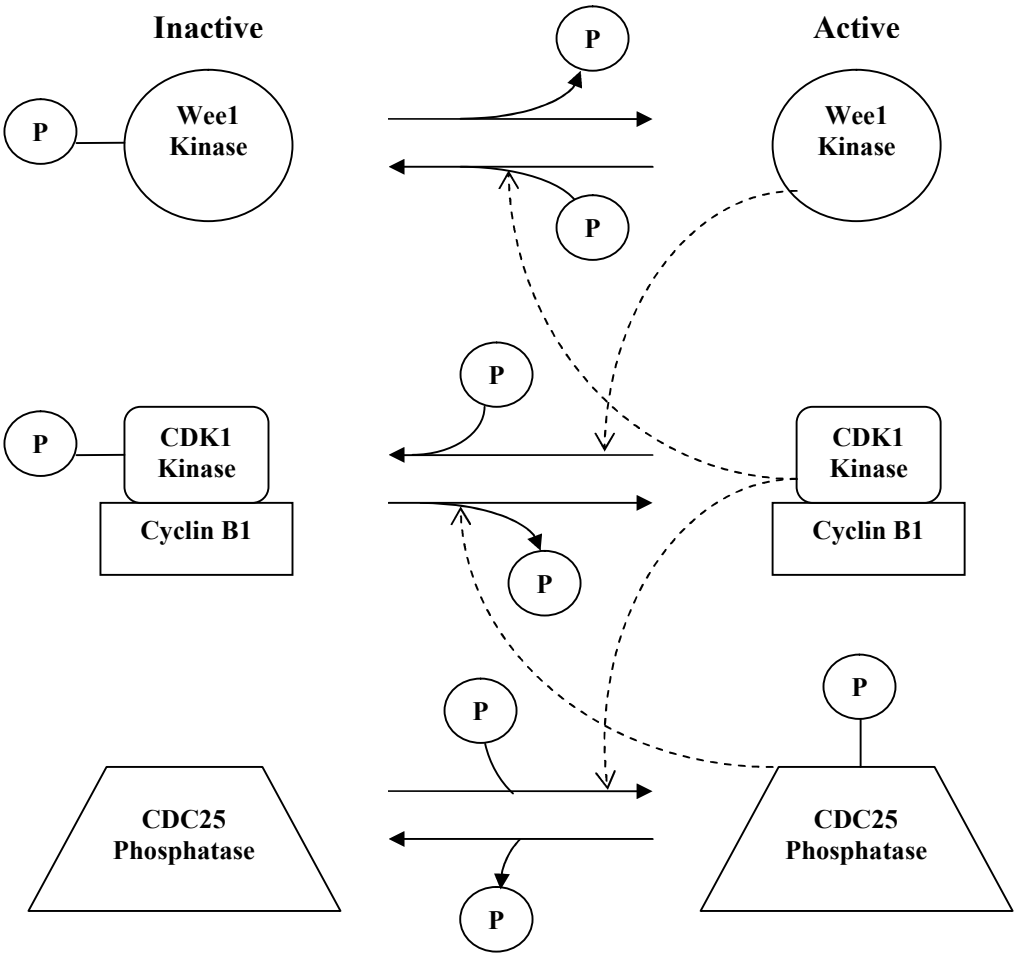


Figure 3.10: Schematic diagram of proposed Wee1 kinase and CDC25 phosphatase interactions with CDK1-Cyclin B1 complexes (redrawn from Reference [156]).

3.4 Summary

This chapter provides a summary of the relevant and necessary biological background in addition to the underlying assumptions that were used in this thesis to develop the extended drug kinetics model, the extended cell cycle model and the coupled drug kinetics/dynamics model.

4 The *in vitro* kinetic model

4.1 Introduction

The state-space modelling approach to drug kinetics can provide a mathematical description of drug stability and delivery to a specified target [15]. For this purpose, a state-space model for the *in vitro* uptake kinetics of the anti-cancer agent TPT has been extended from a previously published model by Evans *et al*. [13] to account for drug resistance mechanisms in cancer cells. That is, the extended model describes the drug activity and delivery of the pharmacologically active form to the nuclear DNA target as well as the catalysis of the ALDH enzyme and the elimination of drug from the cytoplasm via the efflux pump (see Chapter 3). The extended model is based on underlying biological assumptions, and the unknown parameters of the proposed model are estimated from two-photon laser-scanning microscopy (TPLSM) data, which provide concentrations of TPT (lactone plus hydroxy acid forms) in the extracellular region containing live human breast cancer cells (MCF-7 cell line), the cytoplasm and the nucleus.

High performance liquid chromatography (HPLC) provides a means by which the levels (concentrations) of TPT_L and TPT_H can be measured from medium and cellular extracts. Data collected by project collaborators (Wales College of Medicine, Cardiff University) from *in vitro* cell experiments using HPLC provide information about the respective concentrations of TPT_L and TPT_H in the medium and cellular pools. These data have been used to drive model development and to estimate unknown parameters in corresponding state-space models [15].

The reversible hydrolysis (see Section 3.1) of TPT_L was modelled by Evans *et al.* [15] using a simple linear two-compartment model with parameter estimation from HPLC data in buffers at different pH levels in order to characterise the pH dependency of the reversible conversion between TPT_L and TPT_H. According to the results obtained from the live cell experiments, differential reversible hydrolysis of TPT occurs in the medium in

comparison to the cellular pools. This occurs due to the difference between intracellular and extracellular pH levels as well as due to less well-defined properties of the drug in the cell.

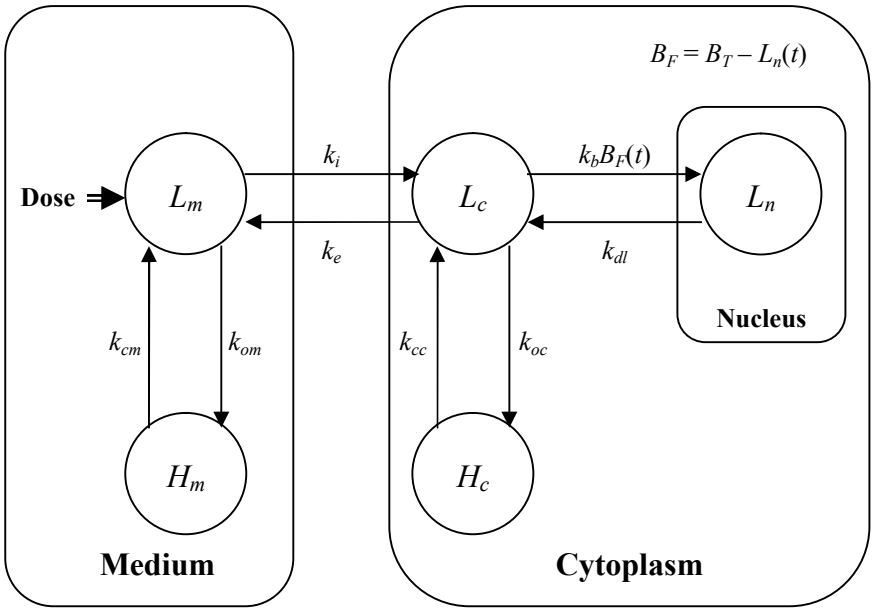


Figure 4.1: Schematic of the five-compartment model (proposed by Evans *et al.* [15]) used to investigate the effect on TPT of injecting a dose (system concentration 10 μM) of TPT_L into a culture medium containing human lymphoma cells (SU-DHL-4 cell line).

Based on the two-compartment model, Evans *et al.* [15] developed a non-linear five-compartment model (Figure 4.1) describing the uptake kinetics of TPT in a medium enclosing human lymphoma cells (SU-DHL-4 cell line). This was achieved by combining compartmental models for both the intracellular and extracellular pools. In addition, cellular influx and efflux terms were included to partially demonstrate the stability of TPT_L in solutions containing live cells. Finally, a compartment was added to represent the cellular target, which is the pharmacologically active form of the drug (TPT_L) bound to nuclear DNA. To complete the mathematical model which describes the drug (TPT_L) activity and delivery to DNA target, appropriate non-linear association and linear dissociation pathways were included. The non-linear model characterises the kinetics of TPT in the entire cell population. Similar to the two-compartment model in Evans *et al.* [15], the unknown parameters for the five-compartment model were estimated from HPLC data. Before estimating the unknown model parameters from the *in vitro* data, a structural identifiability analysis was performed which demonstrated that all of the unknown model

parameters are uniquely determined by the output structure corresponding to the experiment. HPLC allows the direct measurement of the concentrations of TPT_L and TPT_H (separately) in the extracellular medium as well as the concentrations of TPT_L and TPT_H (separately) in the cellular pool (i.e., the cytoplasm and the nucleus).

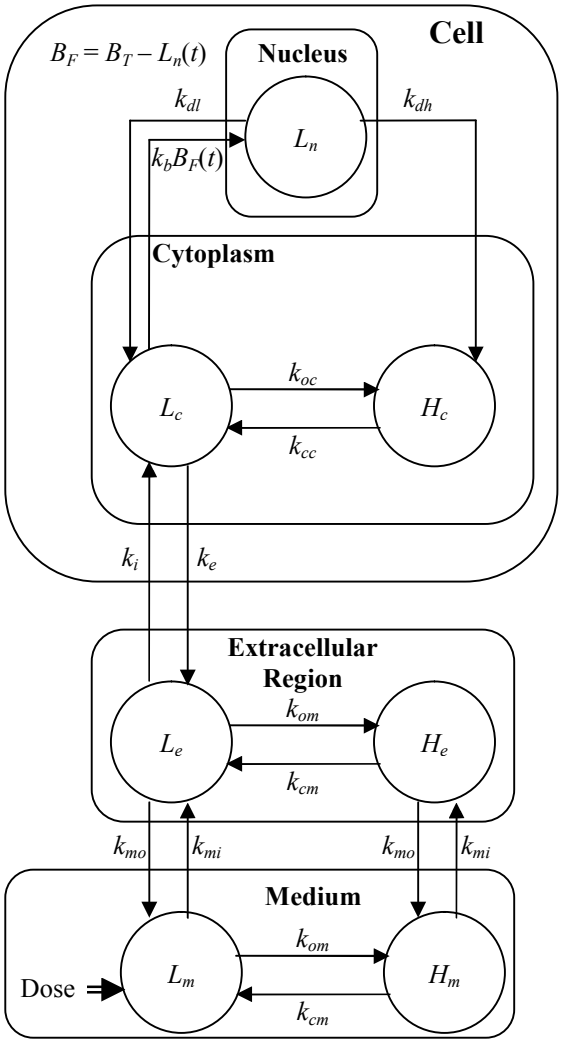


Figure 4.2: Schematic of the mathematical model (developed by Evans *et al.* [13]) used to investigate the uptake kinetics of TPT in a culture medium containing human breast cancer cells (MCF-7 cell line) in suspension.

The five-compartment model (Figure 4.1) does not allow for heterogeneity in the medium; that is, the model does not define the region within the medium pool from which the active form of the drug (i.e., TPT_L) may flow into the cells, since the extracellular compartments are assumed to be homogeneous and well-mixed (by definition) [15]. If the influx of TPT_L to the cell were from a smaller region of the medium (that is not well-mixed with the rest) then the ratio of TPT_L to TPT_H in this region might seem more favoured to TPT_H, however,

this does not match expectations from the pH buffer experiments (see Evans *et al.* [15]). For this reason, the five-compartment model was further developed to a seven-compartment system (Figure 4.2) by Evans *et al.* [13], by splitting the medium into two separate pools (regions): a region into which the drug is added and a region from which drug can flow into the cells (see Section 4.3). Similarly, the seven-compartment model describes the drug activity and delivery of the pharmacologically active form TPT_L to the DNA target. The unknown model parameters (of the seven-compartment model in Figure 4.2) are uniquely identifiable by the output structure corresponding to the experiments and were estimated from two-photon laser-scanning microscopy data from live human breast cancer cells (MCF-7 cell line).

4.2 TPT reversible hydrolysis model

The reversible hydrolysis of TPT is an important aspect of the kinetics of this particular drug and forms the basic building block for the full cell-based model presented in this thesis (Section 4.3). The inactivation of TPT_L and subsequent activation of TPT_H has been modelled by Evans *et al.* [15] using a simple two-compartment model. Figure 4.3 provides a schematic diagram of the two-compartment model developed by Evans *et al.* [15] where $L(t)$ stands for the concentration of parent form of the drug (TPT_L) in solution at time t after an initial dose of TPT_L and $H(t)$ denotes the corresponding concentration of the inactive form of the drug (TPT_H). The ring opening (inactivation) of TPT_L occurs with a first order rate constant k_o to give TPT_H and the first order rate constant k_c is for the activation (ring closing) of TPT_H to give TPT_L. The following coupled linear differential equations describe the dynamics of the concentrations of the active form (TPT_L) and inactive form (TPT_H) of TPT:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{dt}} L(t) &= -k_o L(t) + k_c H(t) \\ \frac{\mathrm{d}}{\mathrm{dt}} H(t) &= k_o L(t) - k_c H(t) \end{aligned} . \tag{4.1}$$

In all of the experiments considered in this thesis, TPT is administered as a single bolus input (impulsive input) of TPT_L (active/lactone form) to yield an initial concentration D for

$L(0)$. Since the drug concentrations are zero before administration (i.e. at $t = 0$), the corresponding initial conditions for the two-compartment model Equation (4.1) are given by:

$$L(0) = D \text{ and } H(0) = 0. \tag{4.2}$$

In Evans *et al.* [15] the values of the unknown model parameters k_o and k_c were estimated from HPLC data collected at Cancer Research Wales Laboratories (Velindre NHS Trust, Cardiff) for different pH values consistent with physiological conditions (pH 6.8-8.0). However, in this thesis, the values for the constants (k_o and k_c) will only be estimated at pH = 7.2 since the MCF-7 cells (cell-based model in Section 4.3) were placed in a medium with pH = 7.2 in the TPLSM experiments. The experiment carried out to collect HPLC data allows separate sampling of the concentrations of TPT_L and TPT_H, accordingly, there are two model outputs (observations), $L(t)$ and $H(t)$. In the pH buffer, concentrations were measured in μM, the initial dose was D = 10 μM. Before estimating the unknown model parameters from the HPLC data, a structural identifiability analysis (using the Laplace transform) for the model was performed.

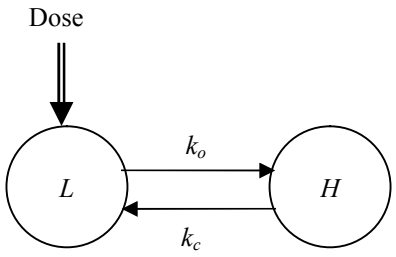


Figure 4.3: Schematic of the simple two-compartment model (developed by Evans *et al.* [15]) used to investigate TPT hydrolysis in buffers at different pH values.

4.2.1 Structural identifiability analysis for the hydrolysis model

The Laplace transform method presented in Section 2.4 is applied to perform the structural identifiability analysis for the two-compartment hydrolysis model developed by Evans *et al*. [15]. The dose D is assumed to be known and is therefore not included in the parameter vector $\mathbf{p}$, which is given by $\mathbf{p} = (k_o, k_c)^T$. Since the parameters must be real and positive,

the set of feasible vectors is taken to be $\Omega = \{(p_1, p_2)^T : p_i > 0 \text{ and real}\}$f. The state vector $\mathbf{x}(t, \mathbf{p})$ ($n = 2$) and output vector $\mathbf{y}(t, \mathbf{p})$ ($r = 2$) are the same and are given by

$$\mathbf{y}(t, \mathbf{p}) = \mathbf{x}(t, \mathbf{p}) = \begin{pmatrix} L(t, \mathbf{p}) \\ H(t, \mathbf{p}) \end{pmatrix}.$$

The state takes values in the connected open subset $M(\mathbf{p})$ which is taken to be $\mathbb{R}^2$ (i.e., the set of all 2-vectors $(L(t), H(t))^T$. Notice that for a unit impulse input (a bolus injection) at $t = 0$ into a zero initial state system, one can write the equations without input as shown in Equations (2.18) and (2.19). Therefore, the initial state $\mathbf{x}(0, \mathbf{p})$ is given by $\mathbf{x}_0(\mathbf{p}) = (D, 0)^T$ which is independent of the vector of unknown model parameters $\mathbf{p}$ (since D is assumed to be known). To obtain the equations similar to the form of Equations (2.18) and (2.19), i.e. non-zero initial state without further input, the following is defined:

$$\begin{pmatrix} \dot{L} \\ \dot{H} \end{pmatrix} = \begin{pmatrix} -k_o & k_c \\ k_o & -k_c \end{pmatrix} \begin{pmatrix} L \\ H \end{pmatrix},$$

where $(L, H)^T$ is an arbitrary vector in $M(\mathbf{p})$, the matrix

$$\mathbf{A}(\mathbf{p}) = \begin{pmatrix} -k_o & k_c \\ k_o & -k_c \end{pmatrix}$$

is the matrix of transfer coefficients which contains the basic (structural) components of the compartmental model, and the output matrix

$$\mathbf{C}(\mathbf{p}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

is independent of the vector of unknown model parameters $\mathbf{p}$.

Taking Laplace transforms of Equations (4.1) and (4.2) and rearranging gives:

$$\begin{aligned} (s + k_o)\mathbf{L}(s) - k_c\mathbf{H}(s) &= D \\ -k_o\mathbf{L}(s) + (s + k_c)\mathbf{H}(s) &= 0 \end{aligned} \tag{4.3}$$

where $\mathbf{L}(s)$ and $\mathbf{H}(s)$ are the Laplace transforms of $L(t)$ and $H(t)$, respectively. Solving for $\mathbf{L}(s)$ and $\mathbf{H}(s)$ results in:

$$\begin{aligned}\mathbf{L}(s) &= \frac{Ds + Dk_c}{s^2 + (k_c + k_o)s} \\ \mathbf{H}(s) &= \frac{Dk_o}{s^2 + (k_c + k_o)s}\end{aligned}. \tag{4.4}$$

Therefore, the Laplace transform of the observation function is:

$$\begin{aligned}\mathbf{Y}(s) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \dfrac{Ds + Dk_c}{s^2 + (k_c + k_o)s} \\ \dfrac{Dk_o}{s^2 + (k_c + k_o)s} \end{pmatrix}, \\ Y_1(s) &= \frac{Ds + Dk_c}{s^2 + (k_c + k_o)s} \\ Y_2(s) &= \frac{Dk_o}{s^2 + (k_c + k_o)s}.\end{aligned} \tag{4.5}$$

The coefficients are the moment invariants (observational parameters), therefore, determining Y_1(s) and Y_2(s) means that values for all of the coefficients are determined. Looking at the coefficients of individual outputs, the equations for the observational parameters in terms the model basic parameters are:

$$\begin{aligned}\phi_1 &= k_o + k_c \\ \phi_2 &= 0 \\ \phi_3 &= D \\ \phi_4 &= Dk_c \\ \phi_5 &= Dk_o\end{aligned}. \tag{4.6}$$

The coefficient $\phi_2 = 0$ as a result of a zero eigenvalue. The coefficient ϕ_4 determines k_c since the value of D is a measurable constant, and from ϕ_1, k_o is determined. Accordingly, the two basic parameters (k_o and k_c) of the model are uniquely determined by the output structure corresponding to the proposed experiment. Therefore, the two-compartment hydrolysis model is structurally globally (uniquely) identifiable. The structural identifiability analysis was performed in this section to estimate the unknown model parameters with greater confidence.

4.2.2 Parameter estimation for the hydrolysis model

Parameter estimation (Appendix B) was performed using the commercial software package FACSIMILE (MCPA Software, U.K.) for HPLC data consisting of values of both forms of TPT (i.e., TPT_L and TPT_H) at times t = 1, 2, 5, 10, 15, 30, 45, 60 and 120 min following drug administration (via a bolus injection). Phosphate buffer saline (PBS) solutions made to the required pH levels (6.8, 7.0, 7.2, 7.4, 7.6, 7.8 and 8.0) were used in the experiments. PBS is an isotonic buffer solution (i.e., the solution has the same ion concentration and osmolarity as the human body) widely used in the field of biological research to maintain constant pH levels during experiments. This buffer solution contains sodium chloride (NaCl), sodium phosphate (Na_2HPO_4) and, in some chemical formulations, it may also include potassium chloride (KCl) and potassium phosphate (KH_2PO_4). Table 4.1 represents data from the HPLC experiments (at pH = 7.2) which were performed in duplicate (i.e., two experiments at each pH level) and Figure 4.4 is the graphical representation of the data in Table 4.1. The unknown parameters (k_o and k_c) for the hydrolysis model were estimated at pH = 7.2 since MCF-7 cells were placed in a medium with pH = 7.2 in the microscopy experiments used to collect data from live cells.

Table 4.1: HPLC data for TPT_L and TPT_H from two different sets (A and B) of TPT hydrolysis experiments at pH = 7.2. Time is given in minutes (min) and concentrations in micromolar (μM).

Time (mins)	Set A: TPT_L (μM)	Set A: TPT_H (μM)	Set B: TPT_L (μM)	Set B:TPTH (μM)
1.00	9.99	0.00	9.93	0.05
2.00	9.89	0.09	9.85	0.11
5.00	9.71	0.27	9.63	0.29
10.00	9.36	0.60	9.26	0.71
15.00	8.98	1.10	8.91	1.02
30.00	7.96	2.01	8.40	1.52
45.00	7.47	2.48	7.34	2.61
60.00	7.10	3.07	6.86	3.08
120.00	6.79	3.11	6.83	3.12

The parameter estimation was performed simultaneously using the two HPLC data sets (four data curves) in Table 4.1. The fitted (estimated) parameter values with estimates for their respective confidence intervals are presented in Table 4.2, where it is seen that the two parameters (k_o and k_c) are well-determined by the HPLC data. In addition, the confidence limits for both parameters (k_o and k_c) are reasonably close, which gives further

confidence that the parameter values are well-determined. In addition, the value of $CORI_{mean} = 0.61467$ indicates that the residuals are acceptable and evenly distributed above and below the (simulated) output of the model and therefore, the resulting fit is good.

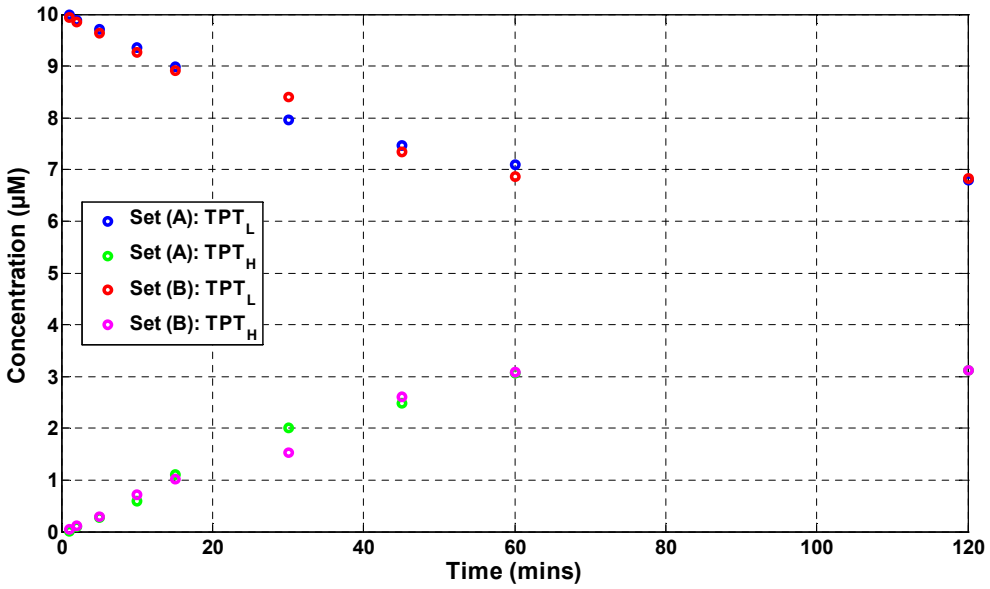


Figure 4.4: Graphical presentation of HPLC data for TPT_L and TPT_H from two different sets (A and B) of TPT hydrolysis experiments at pH = 7.2.

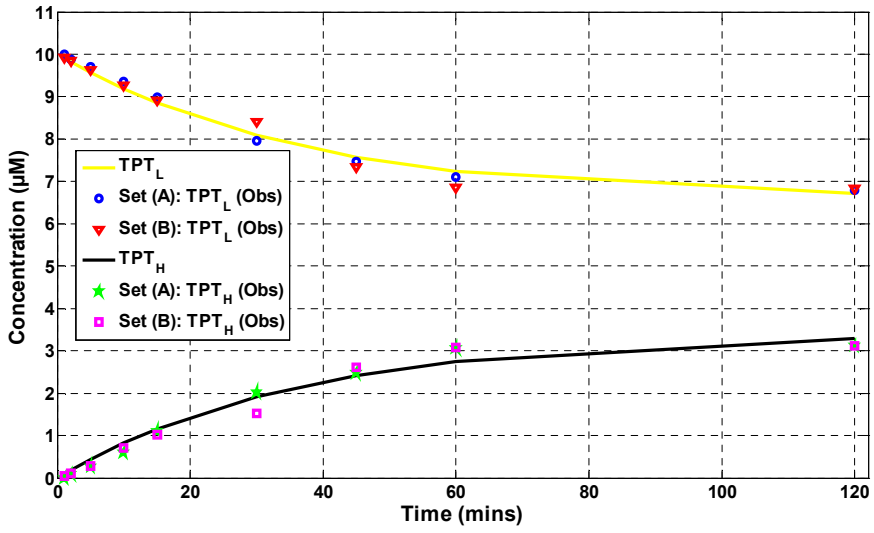


Figure 4.5: The plot represents the model output (lines) for TPT_L (yellow) and TPT_H (black) fitted to the HPLC data sets A (circles and pentagrams for TPT_L and TPT_H respectively) and B (triangles and squares for TPT_L and TPT_H respectively) at pH = 7.2.

Table 4.2: The parameter values k_o and k_c for the model Equations (4.1) and (4.2) obtained from HPLC data (set A and set B).

Parameter	Value	SDLN	5%	95%
k_o (min^{-1})	9.36×10^{-3}	0.0507	8.611×10^{-3}	1.0174×10^{-2}
k_c (min^{-1})	1.8×10^{-2}	0.0935	1.5435×10^{-2}	2.0992×10^{-2}

RSS = $2.7118\times10^{+1}$

In Figure 4.5, the simulated output of the model, given by Equations (4.1)–(4.2), is plotted together with the experimental data. From this plot, it is seen that there is close reproduction of the collected data by the simulated output from the hydrolysis model with parameters taking values from Table 4.2. The correlation coefficient for the two parameters is 0.903 showing high linear covariation between the estimates for $\ln(k_o)$ and $\ln(k_c)$. Accordingly, only a linear combination of the two parameters that is well estimated, and that they are not being estimated separately [23]. The results obtained from the parameter estimation of the unknown parameters of the hydrolysis model, namely, k_o and k_c at pH = 7.2, agree with the parameter estimates reported in Evans *et al.* [15]. The results in Table 4.2 were used in the extended model (cell based model) in the following section (Section 4.3).

From the results obtained by Evans *et al.* [15], it is shown that as the value of pH increases, the long term value of TPT_L decreases and the initial slope of the TPT_L time curve becomes steeper. Thus indicating that at higher pH levels, the initial dose (D) of the active form of the drug (TPT_L) undergoes hydrolysis to the inactive form (TPT_H) more rapidly. In addition, the equilibrium concentration of TPT_L decreases. However, TPT_H is formed more rapidly at higher pH levels with an increased concentration at equilibrium. As expected (see Section 3.1.3), this indicates that TPT_L predominates (i.e., becomes more stable) at lower pH levels. Evans *et al.* [15] have also concluded the estimates for the parameter k_o vary linearly with pH (the estimated value of k_o increases with pH), however, there is not a clear relationship between k_c and pH.

Figure 4.6 is a normal probability plot for the residuals in the weighted non-linear least-squares fit of the hydrolysis model to HPLC data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale. The figure suggests that the residuals are normally distributed

(with approximately zero mean) since the graph is approximately linear. This is an indicator of the appropriateness of the standard error σ_i used for each observation y_i.

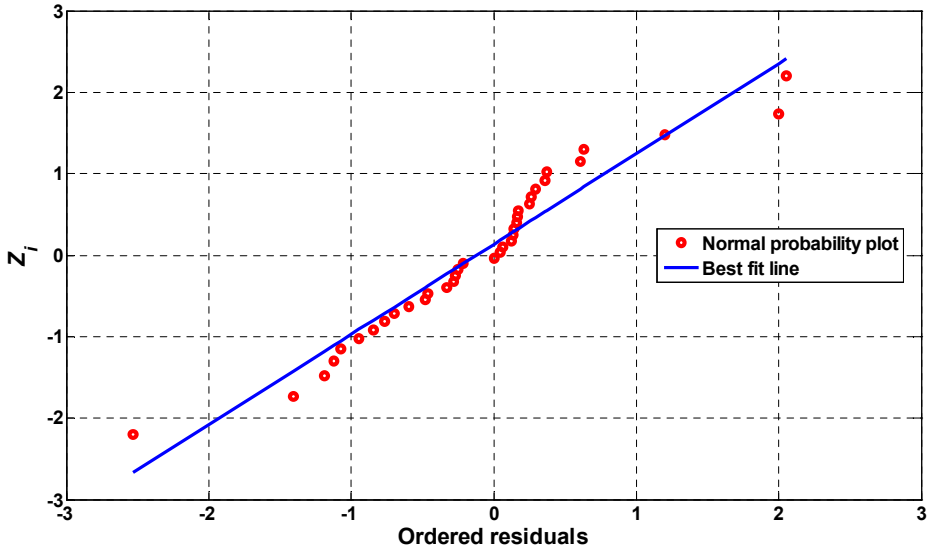


Figure 4.6: A normal probability plot for the residuals in the weighted non-linear least-squares fit of the hydrolysis model to HPLC data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

4.3 The cell-based model

The simple two-compartmental model for the reversible hydrolysis of TPT_L (Figure 4.3) in buffers was used as the main building block to construct the five- and seven-compartment models developed by Evans *et al.* [13, 15] as well as the extended model in this thesis; these models are based on the kinetics of TPT in live cell experiments. Similar to the previous cell-based models, the extended model yields quantitative predictions (through simulations) of the concentrations of both forms of TPT (i.e., TPT_L and TPT_H) in the various experimental pools.

The new single-cell state-space model is a model of processes based on mass balance principles. The model is represented by first order ODEs describing the reactions and transfer of both forms of TPT, within the system. Accordingly, the simulation results

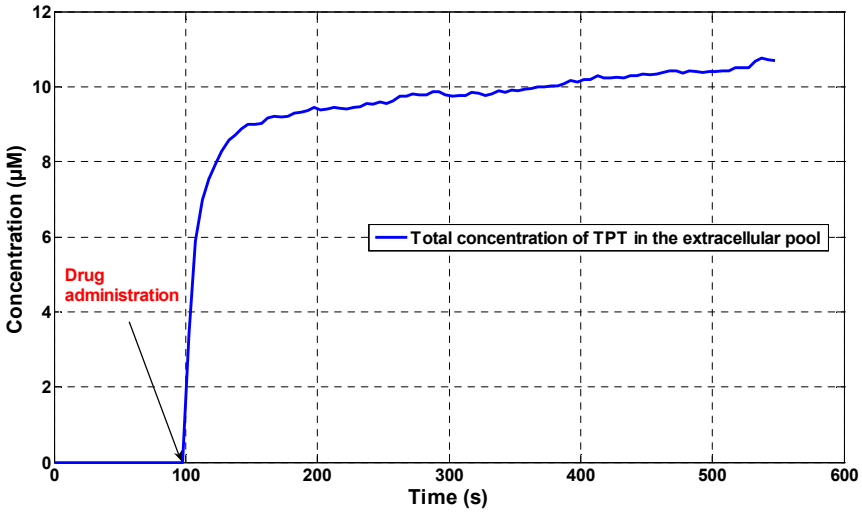
obtained from the model are compared with real experimental data (live human breast cancer cells using two-photon laser-scanning microscopy) for parameter estimation.

4.3.1 Mathematical model

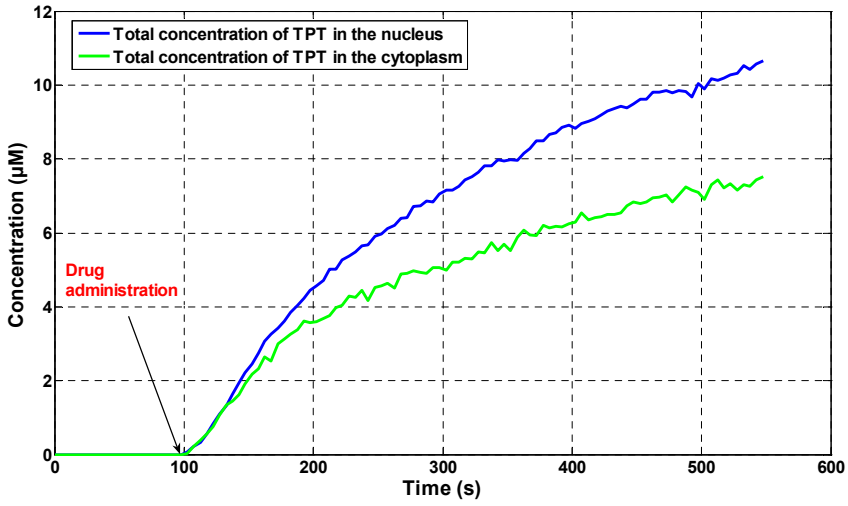
The original single-cell compartmental model (the seven-compartment model in Figure 4.2) proposed by Evans *et al*. [13] allows for heterogeneity by dividing the medium into two separate pools: the medium region into which the drug is added and the extracellular region in which the cells are located. Figure 4.7 demonstrates the response of a single cell for which the data were collected during the two-photon laser-scanning microscopy experiments on MCF-7 cells, and it is evident from Figure 4.7(a) that the mixing takes place in the medium. The drug exchange between the cells and medium takes place only through the extracellular region. In this model, the cellular pool that contains all of the individual cells has been extended to include the cell membrane. Therefore, this leads to a model with five physical regions: the medium, the extracellular region, the cell membrane, the cytoplasm, and the nucleus which are denoted by subscripts m, e, cm, c, and n, respectively. Assumptions about the kinetics of TPT (see Chapter 3) on a single cell basis were made. In addition, the Law of Mass Action was used to model all the chemical reactions within the extended system.

The extended model (Figure 4.8) is made up of nine compartments, each compartment represents a well-mixed homogenous subsystem. These subsystems represent the concentration of the anti-cancer drug in its active (TPT_L) or inactive (TPT_H) forms, in a particular (physical) location with respect to time. The arrows in the model represent transfers between compartments. These flows can be as a result of the reversible hydrolysis of TPT (Section 4.2), the rate of flow of drug (TPT_L or TPT_H) from one physical location to another physical location, the association/dissociation of the drug to/from the drug transporter, the association/dissociation of the drug to/from the ALDH enzyme or the association/dissociation of the drug to/from the nucleus. Figure 4.8 provides a schematic diagram of the extended state-space model, where L and H stand for the concentrations of TPT_L and TPT_H, respectively. Therefore, $L_m(t)$, $L_e(t)$, $L_c(t)$, and $L_n(t)$, denote the concentrations of TPT_L in the medium, extracellular region, cytoplasm and nucleus

respectively, at time t following the addition of drug. The corresponding variables for TPT_H are $H_m(t)$, $H_e(t)$, and $H_c(t)$. The concentration of TPT_H bound to the efflux pump is denoted by $TH_{cm}(t)$, and the variable $T(t)$ represents the total concentration of available ABCG2 (drug transporter) binding sites at time t. TPT_H (the form favoured by the efflux pump) in the cytoplasm (H_c) binds to the drug transporter (T) to form an intermediate complex (TH_{cm}) which breaks in one direction to release free drug transporter (T_0) and TPT_H in the extracellular region (H_e). That is, $H_c + T \xrightarrow{k_1} TH_{cm} \xrightarrow{k_2} T + H_e$.



(a)



(b)

Figure 4.7: The total concentration of TPT (TPT_L + TPT_H) against time in (a) the extracellular pool (b) nucleus (blue line) and cytoplasm (green line). The drug was administered at t = 97.532 seconds.

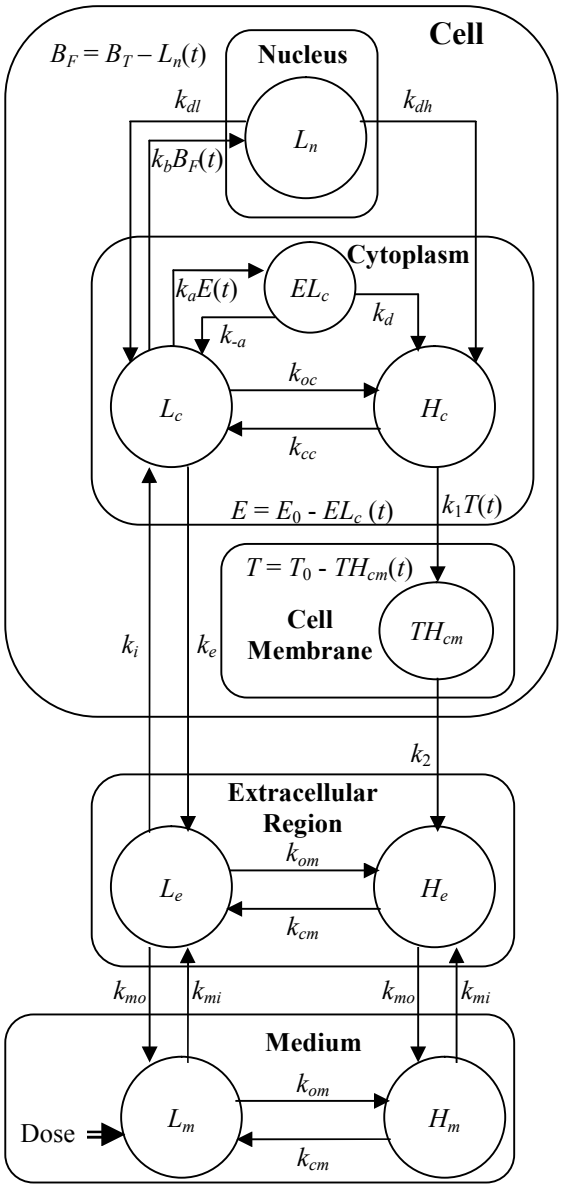


Figure 4.8: Schematic of the extended mathematical model used to investigate the uptake kinetics of TPT in a culture medium containing human breast cancer cells in suspension.

In Evans *et al.* [13, 15] , it is assumed that all drug in the nucleus is bound and only TPT_L binds to DNA. The latter assumption is supported by current knowledge from DNA studies in free solution (see Chapter 3). The former assumption is a simplification of the real case in which the nucleus may contain unbound drug as well; therefore, this postulation requires further experimental investigation. Therefore, in the model, there is no direct pathway for the inactive form of the drug (TPT_H) to enter the nucleus. In the three experimental pools, i.e. medium, extracellular region and cytoplasm, TPT undergoes reversible and pH dependent hydrolysis, as considered in Section 4.2, which has been modelled by a simple two-compartment model [15]. The rate constants for the ring closing of TPT_H and ring opening of TPT_L are identical in the medium and extracellular pools; assuming

homogeneity in pH and/or any other factors that might have an effect on the hydrolysis, the rate constants are k_{cm} and k_{om} for the ring closing of TPT_H and ring opening of TPT_L, respectively, for both the medium and the extracellular region. The corresponding rate constants for the cytoplasm are k_{cc} and k_{oc}.

The active drug (TPT_L) in the cytoplasm is assumed to bind irreversibly (see Chapters 2 and 3) to the enzyme ALDH where it is converted to the form favoured by the efflux pump (TPT_H). In accordance with the Law of Mass Action, the rate at which TPT_L binds to ALDH is assumed to be proportional, with the association rate k_a, to the product of the concentrations of TPT_L in the cytoplasm, $L_c(t)$, and free enzyme, $E(t)$. Additionally, dissociation of drug bound to ALDH occurs at a first-order rate as either TPT_L, with rate constant k_{-a}, or TPT_H, with rate constant k_d. If E_0 denotes the total concentration of available ALDH, then the concentration of free ALDH (by applying suitable conservation laws) is $E(t) = E_0 - EL_c(t)$, where $EL_c(t)$ is the concentration of enzyme-drug complex.

Mixing between the medium and extracellular pools can be modelled by first-order flows between them. Assuming that the rate constants are the same for the lactone and hydroxy acid forms, k_{mi} denotes the flow from the medium to the extracellular pool and k_{mo} denotes the flow out of the extracellular pool to the medium. Flow between the extracellular region and the cellular pool, including the cell membrane and cytoplasm, takes place simultaneously via two distinct processes. The first process is the efflux pumping mechanism (see Chapter 2) in which the BCRP/ABCG2 transporter carries TPT_H through the cell membrane to the extracellular region irreversibly. According to the Law of Mass Action, the rate at which TPT_H in the cytoplasm binds to the efflux pump is assumed to be proportional, with association rate constant k_1, to the product of the concentrations of the inactive drug in the cytoplasm, $H_c(t)$, and free transporter, $T(t)$. The dissociation of drug bound to the efflux pump (TH_{cm}) is assumed to be first order with rate constant k_2. If T_0 denotes the total concentration of available (BCRP/ABCG2) drug transporter, then the concentration of free transporter (by applying suitable conservation laws) is the difference between T_0 and the concentration of bound drug (TH_{cm}), that is $T(t) = T_0 - TH_{cm}(t)$. The second process is the diffusion between the cytoplasm and the extracellular region, which is also modelled via a first-order process. Note however, that only TPT_L diffuses across the cell membrane, based on previous experimental observations [15, 48] and physical

properties of TPT_H with respect to the lipid bilayer traverse [202]. The rate constant for influx of $L_e(t)$ into the cytoplasm is k_i and the rate constant for the efflux of $L_c(t)$ into the extracellular pool is k_e. Similarly, the binding rate of TPT_L to DNA is assumed to be proportional (with constant k_b) to the product of the concentrations of TPT_L in the cytoplasm and free binding sites, $B_F(t)$. Let B_T denote the total concentration of available DNA binding sites, therefore, the concentration of free sites, $B_F(t)$, is the difference between B_T and the concentration of bound drug, $L_n(t)$, that is $B_F(t) = B_T - L_n(t)$. Dissociation of bound drug is assumed to occur at a first-order rate as either TPT_L, with rate constant k_{dl}, or TPT_H, with rate constant k_{dh}. That is, TPT_L may bind to DNA in a reversible manner and can then, once bound, be converted to TPT_H, as suggested by Streltsov [203].

If the volumes of the medium, extracellular region, cytoplasm, cell membrane and nucleus, are denoted by V_m, V_e, V_c, V_{cm}, and V_n, respectively, then the postulated mathematical model for the uptake kinetics of TPT is given by the following system of differential equations:

$$
\begin{aligned}
\frac{\mathrm{d}L_m}{\mathrm{dt}} &= -(k_{om} + k_{mi})L_m + k_{cm}H_m + k_{mo}v_0 L_e \\
\frac{\mathrm{d}H_m}{\mathrm{dt}} &= k_{om}L_m - (k_{cm} + k_{mi})H_m + k_{mo}v_0 H_e \\
\frac{\mathrm{d}L_e}{\mathrm{dt}} &= \frac{k_{mi}L_m}{v_0} - (k_{mo} + k_{om} + k_i)L_e + k_{cm}H_e + \frac{k_e L_c}{v_1} \\
\frac{\mathrm{d}H_e}{\mathrm{dt}} &= \frac{k_{mi}H_m}{v_0} + k_{om}L_e - (k_{cm} + k_{mo})H_e + \frac{v_3 k_2 TH_{cm}}{v_1} \\
\frac{\mathrm{d}L_c}{\mathrm{dt}} &= k_i v_1 L_e - (k_e + k_{oc})L_c + k_{cc}H_c + k_{dl}v_2 L_n \\
&\quad -k_b(B_T - L_n)L_c - k_a(E_0 - EL_c)L_c + k_{-a}EL_c \\
\frac{\mathrm{d}H_c}{\mathrm{dt}} &= k_{oc}L_c - k_{cc}H_c + k_{dh}v_2 L_n - k_1(T_0 - TH_{cm})H_c + k_d EL_c \\
\frac{\mathrm{d}L_n}{\mathrm{dt}} &= \frac{k_b}{v_2}(B_T - L_n)L_c - (k_{dl} + k_{dh})L_n \\
\frac{\mathrm{d}TH_{cm}}{\mathrm{dt}} &= \frac{k_1}{v_3}H_c(T_0 - TH_{cm}) - k_2 TH_{cm} \\
\frac{\mathrm{d}EL_c}{\mathrm{dt}} &= k_a L_c(E_0 - EL_c) - (k_d + k_{-a})EL_c
\end{aligned}
\tag{4.7}
$$

where, $v_0 = V_e/V_m$, $v_1 = V_e/V_c$, $v_2 = V_n/V_c$, $v_3 = V_{cm}/V_c$, and $v_1/v_3 = V_e/V_{cm}$. Time $t = 0$ corresponds to the (first) addition of drug as a bolus injection. The corresponding initial conditions for the model are given by:

$$H_m(0) = L_e(0) = H_e(0) = L_c(0) = H_c(0) = L_n(0) = EL_c(0) = TH_{cm}(0) = 0, \quad (4.8)$$

and

$$L_m(0) = (1 + v_0)D, \quad (4.9)$$

where D is the concentration of dose in the full physical medium (i.e. 2 ml or 2×10^{12} μm^3).

According to Evans *et al.* [13, 15], the estimates obtained for the volumes, using data collected from optical sectioning using a confocal microscope, are: $V_c + V_{cm} = 829$ μm^3 (SD=232 μm^3) for the average volume of cytoplasm and cell membrane in single cell; the volume of the nucleus in a single cell (V_n) is 326 μm^3 (SD=85.5 μm^3); and therefore, the total volume of the average cell ($V_c + V_{cm} + V_n$) is 1155 μm^3. According to the above values, the radius of the cell can be calculated assuming that the cell is spherical, therefore, the average radius of a single cell is 6.5 μm and the average radius of the nucleus in a single cell is 4.27 μm. The cell membrane thickness is approximately 3-10 nm [204]. There is therefore relatively little variation in the plasma membrane thickness between cells. In this model, the chosen value of the membrane thickness is 7 nm. The resulting calculations for the average volumes of V_{cm}, and V_c in a single cell are 8.4 μm^3 and 820.5 μm^3, respectively. The culture medium has a volume of 2×10^{12} μm^3, therefore:

$$V_m + V_e = 2\times10^{12} \text{ μm}^3,$$

and so,

$$V_e = \frac{2\times10^{12} v_0}{1+v_0} \text{ μm}^3.$$

Given that the culture medium contains 1×10^5 cells in the experiment providing data, the cellular volume ratios are given by:

$$v_1 = \frac{V_e}{V_c} = \left(\frac{v_0}{1+v_0}\right)\left(\frac{2\times10^{12}}{(1\times10^5)\times820.5}\right) = \frac{\alpha v_0}{1+v_0}$$

where $\alpha = 2.4375\times10^4$, $v_2 = 3.9732\times10^{-1}$, and $v_3 = 1.0238\times10^{-2}$. Cells were placed in a medium with pH = 7.2 in the experiments used to collect data for estimating the unknown

parameters in the proposed model. With the assumption that the hydrolysis rate constants are primarily dependent on pH, the corresponding values for k_{om}, k_{cm}, k_{oc} and k_{cc} from Section 4.2 (i.e., k_o and k_c) are used, these values are $k_{om} = k_{oc} = 1.6\times10^{-4}$ s^{-1} (9.36×10^{-3} min^{-1}) and $k_{cm} = k_{cc} = 3.0\times10^{-4}$ s^{-1} (1.8×10^{-2} min^{-1}). In the experiment used to collect data for parameter estimation, concentrations were measured in μM, the initial dose was $D = 10$ μM, and time was measured in seconds.

4.3.2 Pseudo-steady state approximation of the cell-based model

In this subsection, the extended model in Figure 4.8 is reduced by assuming that the binding kinetics of both the ALDH enzyme and the BCRP/ABCG2 efflux pump are saturable within the physical volume of space considered. That is, the concentrations of the enzyme and the transporter are much lower than the concentration of associated substrates. A structural identifiability analysis (Section 4.4) of the unknown parameters of the system is performed for the reduced system following a PSSA (see Section 2.3.2) [27]. The structural identifiability properties of the reduced model will change from those of the original system [205]. Therefore, fewer system parameters (of the basic model) may be structurally globally identifiable as they enter in new parameter groupings in the reduced order model. The assumptions on the relative sizes of the model parameters, are the following:

- The concentration of the enzyme is much lower than the active form of the drug in the cytoplasm, that is $E_0 << L_c$ and/or the binding affinity of the enzyme ALDH is much lower than unity, $(k_d + k_{-a})/k_a << 1$.
- The concentration of the transporter (BCRP/ABCG2) is much lower than the concentration of the inactive form of the drug in the cytoplasm, $T_0 << H_c$ and/or the binding affinity of the transporter is low $k_2/k_1 << 1$.

Under these assumptions, following the administration of drug, saturation of the available binding sites (enzyme and transporter) or equilibrium of bound substances is virtually instantaneous, whereby a PSSA can be made [205]. Such an approximation yields a coupled set of algebraic and differential equations. According to the assumptions above, the intermediate compounds TH_{cm} and EL_c are formed rapidly and reach an equilibrium

state (i.e., $\frac{\mathrm{d}TH_{cm}}{\mathrm{dt}} = 0$ and $\frac{\mathrm{d}EL_c}{\mathrm{dt}} = 0$). Therefore, solving for TH_{cm} and EL_c in Equation (4.7) results in:

$$TH_{cm} = \frac{T_0 k_1 H_c}{k_2 v_3 + k_1 H_c} \text{ and } EL_c = \frac{E_0 k_a L_c}{(k_d + k_{-a}) + k_a L_c}. \tag{4.10}$$

Using the above results and substituting these in the state space system (see Equation (4.7)), the ODEs for H_e, L_c and H_c can be reduced to the following:

$$\begin{aligned}
\frac{\mathrm{d}H_e}{\mathrm{dt}} &= \frac{k_{mi}}{v_0} H_m + k_{om} L_e - (k_{cm} + k_{mo}) H_e + \frac{v_3 V_{m1} H_c}{v_1 (v_3 k_{m1} + H_c)} \\
\frac{\mathrm{d}L_c}{\mathrm{dt}} &= k_i v_1 L_e - (k_e + k_{oc}) L_c + k_{cc} H_c + k_{dl} v_2 L_n - k_b (B_T - L_n) L_c - \frac{V_{\mathrm{m}2} L_c}{(k_{m2} + L_c)} \\
\frac{\mathrm{d}H_c}{\mathrm{dt}} &= k_{oc} L_c - k_{cc} H_c + k_{dh} v_2 L_n - \frac{v_3 V_{m1} H_c}{(v_3 k_{m1} + H_c)} + \frac{V_{m2} L_c}{(k_{m2} + L_c)}
\end{aligned} \tag{4.11}$$

where $k_{m1} = k_2/k_1$, $k_{m2} = (k_d + k_{-a})/k_a$, $V_{m1} = k_2 T_0$ and $V_{m2} = k_d E_0$. The k_{mr} values ($r = 1, 2$) represent the relevant Michaelis-Menten constants and the V_{mr} values ($r = 1, 2$) are the numerical constants [205] that represent the relevant maximum velocities obtained when the enzyme and the transporter exist completely in the form TH_c ($r = 1$) and EL_c ($r = 2$), respectively. Such (outer solution) approximations provide accurate results for the asymptotic (long-time) behaviour of drug uptake [205]. Moreover, the number of unknown parameters in the model (Figure 4.9) has been reduced by two and the system equations are reduced to seven. For completeness, the reduced mathematical model is given by:

$$\begin{aligned}
\frac{\mathrm{d}L_m}{\mathrm{dt}} &= -(k_{om} + k_{mi}) L_m + k_{cm} H_m + k_{mo} v_0 L_e \\
\frac{\mathrm{d}H_m}{\mathrm{dt}} &= k_{om} L_m - (k_{cm} + k_{mi}) H_m + k_{mo} v_0 H_e \\
\frac{\mathrm{d}L_e}{\mathrm{dt}} &= \frac{k_{mi} L_m}{v_0} - (k_{mo} + k_{om} + k_i) L_e + k_{cm} H_e + \frac{k_e L_c}{v_1} \\
\frac{\mathrm{d}H_e}{\mathrm{dt}} &= \frac{k_{mi}}{v_0} H_m + k_{om} L_e - (k_{cm} + k_{mo}) H_e + \frac{v_3 V_{m1} H_c}{v_1 (v_3 k_{m1} + H_c)} \\
\frac{\mathrm{d}L_c}{\mathrm{dt}} &= k_i v_1 L_e - (k_e + k_{oc}) L_c + k_{cc} H_c + k_{dl} v_2 L_n - k_b (B_T - L_n) L_c - \frac{V_{\mathrm{m}2} L_c}{(k_{m2} + L_c)} \\
\frac{\mathrm{d}H_c}{\mathrm{dt}} &= k_{oc} L_c - k_{cc} H_c + k_{dh} v_2 L_n - \frac{v_3 V_{m1} H_c}{(v_3 k_{m1} + H_c)} + \frac{V_{m2} L_c}{(k_{m2} + L_c)} \\
\frac{\mathrm{d}L_n}{\mathrm{dt}} &= \frac{k_b}{v_2} (B_T - L_n) L_c - (k_{dl} + k_{dh}) L_n
\end{aligned} \tag{4.12}$$

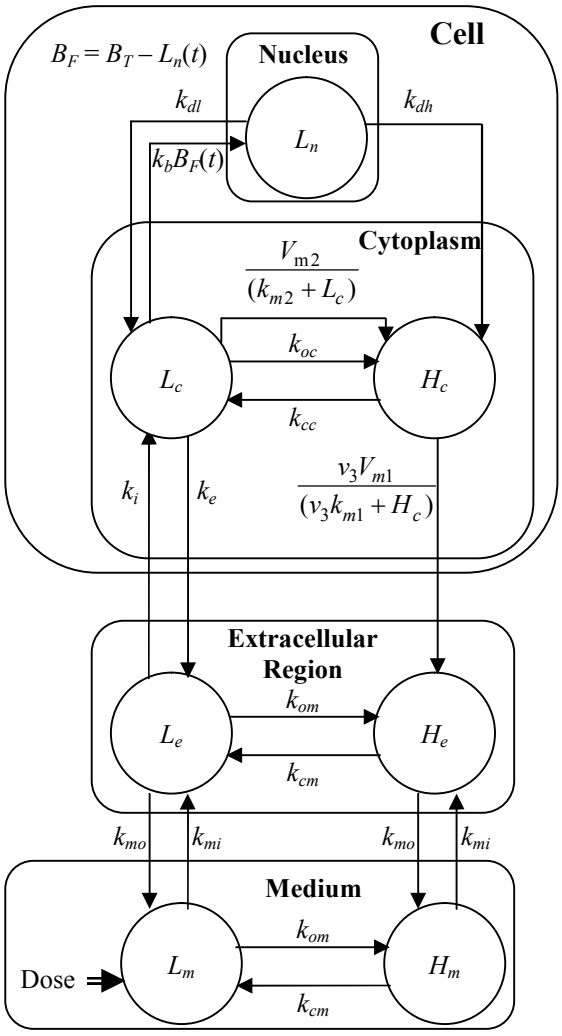


Figure 4.9: Schematic of the reduced mathematical model used to investigate the uptake kinetics of TPT in a culture medium containing human breast cancer cells (MCF-7 cell line) in suspension.

In all of the two-photon laser-scanning microscopy experiments (from the MCF-7 cells) considered in this chapter, concentrations are measured in μM and time t in seconds. For completeness, the twenty parameters in the drug kinetics model (in Figure 4.9) are:

- k_{om} and k_{cm}, units: s^{-1} (per second). Rate constants for deactivation/activation of TPT_L/ TPT_H, respectively, in the medium and extracellular region;
- k_{mi} and k_{mo}, units: s^{-1}. Rate constants for the flow of TPT_L and TPT_H to/from the extracellular region/medium respectively;
- k_i and k_e, units: s^{-1}. Rate constants for the flow of TPT_L into and out of the cell (i.e., to/from extracellular region), respectively;
- k_{oc} and k_{cc}, units: s^{-1} (per second). Rate constants for deactivation/activation of TPT_L/ TPT_H, respectively, in the cell (cytoplasm);

- k_b, units: $(\mu M\ s)^{-1}$. Rate constant for binding (association) of TPT_L (in the cytoplasm) to DNA;
- k_{dl} and k_{dh}, units: s^{-1}. Rate constants of dissociation of TPT_L from DNA to TPT_L/ TPT_H (in cytoplasm) respectively;
- B_T units: μM. Concentration of DNA sites available for TPT_L (in cytoplasm) to bind to;
- k_{m1} and k_{m2}, units: μM. Michaelis-Menten constants for the BCRP/ABCG2 drug transporter and the ALDH enzyme respectively.
- V_{m1} and V_{m2}, units: $\mu M\ s^{-1}$. Numerical constants that represent the maximum velocity obtained when the transporter and the enzyme exist completely in the form TH_c and EL_c respectively;
- v_0, v_1, v_2 and v_3. The ratios, respectively, of the volumes of extracellular region (V_e) to medium (V_m), extracellular region to cytoplasm (V_c), nucleus (V_n) to cytoplasm and cell membrane (V_{cm}) to cytoplasm.

Thirteen of these parameters (k_i, k_e, k_b, k_{mi}, k_{mo}, k_{dl}, k_{dh}, v_0, B_T, V_{m1}, V_{m2}, k_{m1} and k_{m2}) are unknown and need to be estimated from experimental data. Using TPLSM, it is possible to directly measure the total concentration of drug within the extracellular region ($L_e + H_e$), cytoplasm ($L_c + H_c$) and nucleus (L_n). The cell membrane (plasma membrane) acts as a barrier and the drug is transported either passively (TPT_L) or actively (TPT_H) through it. Accordingly, there are no measurements obtained from it (i.e., from the cell membrane as a physical region). Prior to the parameter estimation for the extended drug kinetics model (presented in this chapter), a structural identifiability analysis has been performed (Section 4.4) to test whether the unknown model parameters are uniquely determined by the output structure corresponding to the microscopy experiments. In the following section, the Taylor series approach is used to show that the extended single-cell model (Figure 4.9) is structurally globally identifiable.

4.4 Structural identifiability analysis of the single-cell model

In this section, the Taylor series approach (see Section 2.4) is applied to the pseudo-steady state model (Figure 4.9) described by Equations (4.12) with initial conditions given in Equations (4.8) and (4.9). The input in this particular experiment is a bolus injection of the anti-cancer drug TPT into the system. Therefore, the corresponding amounts can be included as non-zero initial conditions $\mathbf{x}_0(\mathbf{p})$ for the system with no input. The vector $\mathbf{p}$ comprising the thirteen unknown parameters of the model is given by:

$$\mathbf{p} = (k_{mi}, k_{mo}, k_i, k_e, k_b, k_{dl}, k_{dh}, B_T, v_0, V_{m1}, k_{m1}, V_{m2}, k_{m2})^{\mathsf{T}}.$$

The unknown model parameters represent flow rate constants (k_i, k_e, k_b, k_{mi}, k_{mo}, k_{dl} and k_{dh}), Michaelis-Menten constants (k_{m1} and k_{m2}), numerical constants (V_{m1} and V_{m2}), volume ratio (v_0) or a concentration (B_T) and are therefore, positive. This means that the set of feasible parameter vectors, Ω, comprises the vectors $(p_1, \ldots, p_{13})^{\mathsf{T}}$ such that $p_i > 0$ $(1 \leq i \leq 13)$. The state space vector $\mathbf{x}(t, \mathbf{p})$ is given by:

$$\mathbf{x}(t, \mathbf{p}) = (L_m(t, \mathbf{p}), H_m(t, \mathbf{p}), L_e(t, \mathbf{p}), H_e(t, \mathbf{p}), L_c(t, \mathbf{p}), H_c(t, \mathbf{p}), L_n(t, \mathbf{p}))^{\mathsf{T}} \quad (4.13)$$

and the set $M(\mathbf{p})$ is $\mathbb{R}^7$. The initial conditions for the state-space representation of the model (for D = 10 μM) are given by:

$$\mathbf{x}_0(t, \mathbf{p}) = ((1 + v_0)D, 0, 0, 0, 0, 0, 0)^{\mathsf{T}}.$$

The output function of the extended model is given by:

$$\mathbf{y}(t, \mathbf{p}) = (L_e(t, \mathbf{p}) + H_e(t, \mathbf{p}), L_c(t, \mathbf{p}) + H_c(t, \mathbf{p}), L_n(t, \mathbf{p}))^{\mathsf{T}}$$

thus, in terms of the general mathematical formulation of Equation (2.19), the output matrix $\mathbf{C}(\mathbf{p})$ is given by:

$$\mathbf{C(p)} = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (4.14)$$

An arbitrary parameter vector is denoted by $\bar{\mathbf{p}}$ such that:

$$\bar{\mathbf{p}} = (\bar{k}_{mi}, \bar{k}_{mo}, \bar{k}_i, \bar{k}_e, \bar{k}_b, \bar{k}_{dl}, \bar{k}_{dh}, \bar{B}_T, \bar{v}_0, \bar{V}_{m1}, \bar{k}_{m1}, \bar{V}_{m2}, \bar{k}_{m2})^{\mathsf{T}}$$

for which, $\mathbf{y}(t, \mathbf{p}) = \mathbf{y}(t, \bar{\mathbf{p}})$ for all $t \geq 0$. In order to perform the Taylor Series method (Section 2.4), successive coefficients of the Taylor series expansions of the components of $\mathbf{y}(t, \mathbf{p})$ and $\mathbf{y}(t, \bar{\mathbf{p}})$ are computed to determine the relationships that exist between $\mathbf{p}$ and $\bar{\mathbf{p}}$. Computer algebra systems such as MATHEMATICA [206] have proven helpful for the symbolic calculation of these coefficients, particularly when they become algebraically complicated. This software tool was therefore used to perform the following analysis (see Appendix A).

As explained in Subsection 2.3.3, uniqueness of the coefficients in the Taylor series expansions of $y_i(t, \mathbf{p})$ entails the following, if $\bar{\mathbf{p}} \in \boldsymbol{\Omega}$ is such that $\mathbf{y}(t, \mathbf{p}) = \mathbf{y}(t, \bar{\mathbf{p}})$ for all $t \geq 0$, then, for each i = 1, 2, and 3 (for this particular model output), and k = 1, 2, 3, …, then

$$y_i^{(k)}(0,\mathbf{p}) = y_i^{(k)}(0,\bar{\mathbf{p}}) \text{ where, } y_i^{(0)}(0, \mathbf{p}) = y_i(0, \mathbf{p}). \tag{4.15}$$

For k = 0 in Equation (4.15) for i = 1, 2 and 3, no information is gained since each of these coefficients is 0, i.e., $y_i^{(k)}(0,\mathbf{p})$ = 0 for each i = 1, 2, and 3, and k = 0. Now considering the 1st derivative (k = 1) yields:

$$\mathbf{y}^{(1)}(0,\mathbf{p}) = \left(\frac{Dk_{mi}(1+v_0)}{v_0}, 0, 0\right)^{\mathrm{T}} \tag{4.16}$$

and substituting in Equation (4.15) gives:

$$\bar{k}_{mi} = \frac{(k_{mi}(1+v_0)\bar{v}_0)}{((1+\bar{v}_0)v_0)}. \tag{4.17}$$

The 2nd derivative terms (k = 2) give:

$$\mathbf{y}^{(2)}(0,\mathbf{p}) = \left(\frac{-Dk_{mi}(k_i + k_{mi} + k_{mo})(1+v_0)}{v_0}, Dk_i k_{mi}\alpha, 0\right)^{\mathrm{T}} \tag{4.18}$$

and using the relation in Equation (4.17) yields:

$$\bar{k}_{mo} = \frac{k_i(\bar{v}_0 - v_0)}{\bar{v}_0(1+v_0)} + \frac{v_0(k_{mi} + k_{mo}) + \bar{v}_0(k_{mo}v_0 - k_{mi})}{v_0(1+\bar{v}_0)} \text{ and } \bar{k}_i = \frac{k_i v_0(1+\bar{v}_0)}{(1+v_0)\bar{v}_0}. \tag{4.19}$$

Now considering the next derivative (k = 3) in addition to using the relations in Equations (4.17) and (4.19) it follows that:

$$\bar{B}_T = \frac{B_T k_b}{\bar{k}_b}, \bar{k}_e = k_e \text{ and } \bar{v}_0 = v_0. \tag{4.20}$$

Considering the relations (Equations (4.17), (4.19) and (4.20)) between parameters in **p** and $\bar{\mathbf{p}}$, the following equation must hold to satisfy Equation (4.15) for each *i* (1, 2 and 3) and *k* = 0, …, 3:

$$\bar{B}_T = \frac{B_T k_b}{\bar{k}_b}, \bar{k}_e = k_e, \bar{k}_i = k_i, \bar{k}_{mi} = k_{mi}, \bar{k}_{mo} = k_{mo} \text{ and } \bar{v}_0 = v_0. \tag{4.21}$$

Calculating the fourth derivative terms (*k* = 4) and using the relations in Equation (4.21), yields:

$$\begin{aligned} &\bar{k}_{dl} = k_{dh} - \bar{k}_{dh} + k_{dl} + \frac{V_{m2}}{k_{m2}} - \frac{\bar{V}_{m2}}{\bar{k}_{m2}} \text{ and} \\ &\bar{V}_{m1} = \frac{\bar{k}_{m1} k_{m2} \bar{k}_{m2} k_{oc} V_{m1} - k_e k_{m1} \bar{k}_{m1} \bar{k}_{m2} V_{m2} + \bar{k}_{m1} \bar{k}_{m2} V_{m1} V_{m2} + k_e k_{m1} \bar{k}_{m1} k_{m2} \bar{V}_{m2}}{k_{m1} k_{m2} \bar{k}_{m2} k_{oc} + k_{m1} k_{m2} \bar{V}_{m2}}. \end{aligned} \tag{4.22}$$

From the next set of coefficients in the Taylor series expansions (*k* = 5), as well as using the relations in Equations (4.21) and (4.22), gives:

$$\bar{k}_{dh} = k_{dh} \text{ and } \bar{V}_{m2} = \frac{\bar{k}_{m2} V_{m2}}{k_{m2}}. \tag{4.23}$$

Considering the relations (Equations (4.21)-(4.23)) between parameters in **p** and $\bar{\mathbf{p}}$, the following relations must hold in order to satisfy Equation (4.15) for each *i* (1, 2 and 3) and *k* = 0, …, 5:

$$\begin{aligned} &\bar{B}_T = \frac{B_T k_b}{\bar{k}_b}, \bar{k}_{dh} = k_{dh}, \bar{k}_{dl} = k_{dl}, \bar{k}_e = k_e, \bar{k}_i = k_i, \bar{k}_{mi} = k_{mi}, \\ &\bar{k}_{mo} = k_{mo}, \bar{v}_0 = v_0, \bar{V}_{m1} = \frac{\bar{k}_{m1} V_{m1}}{k_{m1}} \text{ and } \bar{V}_{m2} = \frac{\bar{k}_{m2} V_{m2}}{k_{m2}}. \end{aligned} \tag{4.24}$$

Equating and substituting the parameter relations resulting from evaluating the sixth derivative (*k* = 6) in Equation (4.24) gives:

$$\bar{k}_b = k_b, \text{ and } \bar{k}_{m2} = k_{m2}. \tag{4.25}$$

Considering the relations in Equations (4.24) and (4.25) between parameters in **p** and $\bar{\mathbf{p}}$, the following relations must hold in order to satisfy Equation (4.15) for each *i* (1, 2 and 3) and *k* = 0, …, 6:

$$\begin{aligned}&\overline{B}_T = B_T, \overline{k}_b = k_b, \overline{k}_{dh} = k_{dh}, \overline{k}_{dl} = k_{dl}, \overline{k}_e = k_e, \overline{k}_i = k_i, \overline{k}_{m2} = k_{m2}\\ &\overline{k}_{mi} = k_{mi}, \overline{k}_{mo} = k_{mo}, \overline{v}_0 = v_0, \overline{V}_{m1} = \frac{\overline{k}_{m1} V_{m1}}{k_{m1}} \text{ and } \overline{V}_{m2} = V_{m2}.\end{aligned} \tag{4.26}$$

Finally, equating the eighth set of coefficients (k = 7), and combining with the relations in Equation (4.26), yields:

$$\overline{k}_{m1} = k_{m1}. \tag{4.27}$$

Therefore, the following relations must hold in order to satisfy Equation (4.15) for each i (1, 2 and 3) and k = 0, …, 7:

$$\begin{aligned}&\overline{B}_T = B_T, \overline{k}_b = k_b, \overline{k}_{dh} = k_{dh}, \overline{k}_{dl} = k_{dl}, \overline{k}_e = k_e, \overline{k}_i = k_i, \overline{k}_{m2} = k_{m2},\\ &\overline{k}_{m1} = k_{m1}, \overline{k}_{mi} = k_{mi}, \overline{k}_{mo} = k_{mo}, \overline{v}_0 = v_0, \overline{V}_{m2} = V_{m2} \text{ and } \overline{V}_{m1} = V_{m1}.\end{aligned} \tag{4.28}$$

Therefore, for Equation (4.15) to be satisfied $\mathbf{p} = \overline{\mathbf{p}}$ (for each i and all k). This is true for generic $\mathbf{p} \in \Omega$, and thus the model is structurally globally identifiable, that is all of the model parameters are uniquely determined by the output structure corresponding to the proposed experiments to collect data for the purpose of parameter estimation. In the next section, the unknown model parameters were estimated from TPLSM data. The structural identifiability analysis was performed in this section to approach the estimation of the unknown model parameters with greater confidence.

4.5 Parameter estimation of the single-cell model

4.5.1 Two-photon laser-scanning microscopy experiments

Experimental data used for estimating unknown parameters of the extended single-cell model were collected (by the project collaborators at Cardiff University) from an *in vitro* study on the interaction between the anti-cancer drug TPT and live human breast cancer cells (MCF-7 cell line). In the experiments conducted at Cardiff, delivery and localisation of the drug in single MCF-7 cells were examined using TPLSM. The set up of the instrument was calibrated in order that the fluorescence response was linear for 0-15 μM

TPT, making the conversion to the concentration of drug from fluorescence intensity (following background subtraction) simple to calculate. The optical sectioning capacity of this technique provides high-resolution spatial information representing the fluorescence intensity in the compartments and, therefore, quantification of the drug concentration in these compartments, i.e. nucleus (L_e), cytoplasm ($L_c + H_c$) and extracellular ($L_e + H_e$) environment, as shown in Figure 4.10. The kinetics of the uptake for each sub-compartment can be analysed via time-lapse sequences thus providing the primary data for parameter estimation. During the course of the entire uptake assay the signal-to-noise ratio ranged from 1 to 10, this is considered a low value for high resolution microscopy [13]. However, this range for physiological imaging is relatively adequate.

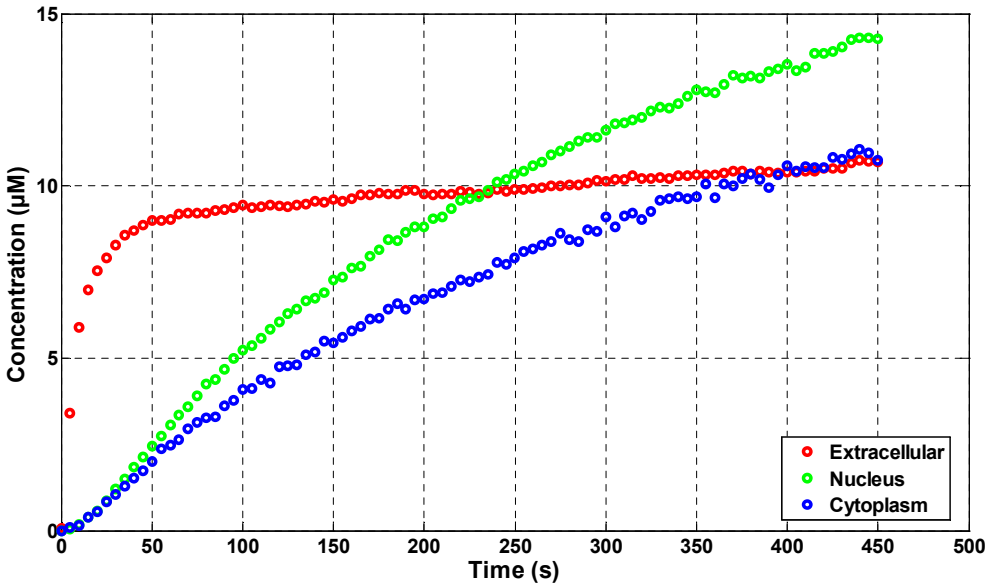


Figure 4.10: Experimental data from TPLSM experiments for total concentration of TPT in the extracellular region (red circles), nucleus (green circles) and cytoplasm (blue circles).

In the TPLSM experiment [12] to track the delivery to sub-cellular compartments of a population of live human breast cancer cells, a concentration of 10 μM of TPT_L was added to a culture of 10^5 adherent cells. Measuring all 10^5 MCF-7 cells was not feasible and therefore 104 cells were physically measured. From these, 13 individual cells were selected as being representative of the entire population in terms of heterogeneity in response to TPT. The duration of the experiment was 7.5 min (450 seconds) with observations (measurements) taken every 5 seconds (Figure 4.10). The dynamic fluorescent intensities representing the drug concentration were derived from the three main cell compartments:

the nucleus, cytoplasm and extracellular region. Experimental data points were collected starting from $t_1 = 5$ seconds to $t_{90} = 450$ seconds, resulting in a data set of 27 observations at 90 time points, that is, 1 mutual measurement for all cells from the extracellular region ($L_e + H_e$) and 2 measurements per cell (for 13 individual cells) from the cytoplasm ($L_c + H_c$) and the nucleus L_n (Figure 4.10).

4.5.2 Parameter estimation

Parameter estimation (Appendix B) was performed using the commercial simulation software package FACSIMILE (MCPA Software, U.K.). The optimisation method used to obtain parameter estimates involves the minimisation of the weighted least-squares criterion (see Section 2.5).

Assuming that TPT hydrolysis is primarily dependent on the level of pH and that there are no other cellular factors affecting it, the hydrolysis rate constants in the cytoplasm (k_{oc} and k_{cc}) were fixed at values estimated from HPLC data at pH = 7.2, (1.6×10^{-4} s^{-1} and 3.0×10^{-4} s^{-1}, respectively), see Section 4.2. The structural global identifiability of the extended model demonstrates uniqueness of the model output for given cellular TPT hydrolysis rate constants, with respect to the unknown parameters.

The response to the anti-cancer drug TPT varies between the 13 individual cells. Therefore, the impact of a dose D can be measured in terms of the total concentration of TPT_L binding to DNA (L_n) over the first 7.5 minutes (450 seconds) following the drug administration which can be represented as the area under the time series curve (AUC) for L_n corresponding to the initial dose D (AUC(L_n, D)). Figure 4.11 shows the measurements (observations) from the nucleus for the 13 individual cells. With reference to Figure 4.11, clearly the responses vary across cells. Therefore, to account for this variation, three different types of curve fitting were conducted for the model represented in Equations (4.12) (at pH = 7.2):

- The data from the 13 individual cells were averaged and the postulated model fitted to these averaged data.

- The model was fitted to data from an individual high loading cell (relatively high concentration of $L_n(t)$).
- Data from an individual low loading cell (relatively low concentration of $L_n(t)$) were used in the fitting.

Table 4.3: *Normalised* AUC(L_n, D) for L_n following a bolus injection D = 10 µM

Cell	AUC
Cell 1	0.46148
Cell 2	0.89843
Cell 3	0.79741
Cell 4	0.44891
Cell 5	0.69542
Cell 6	0.98369
Cell 7	1
Cell 8	0.94444
Cell 9	0.50931
Cell 10	0.65691
Cell 11	0.79689
Cell 12	0.85242
Cell 13	0.86333

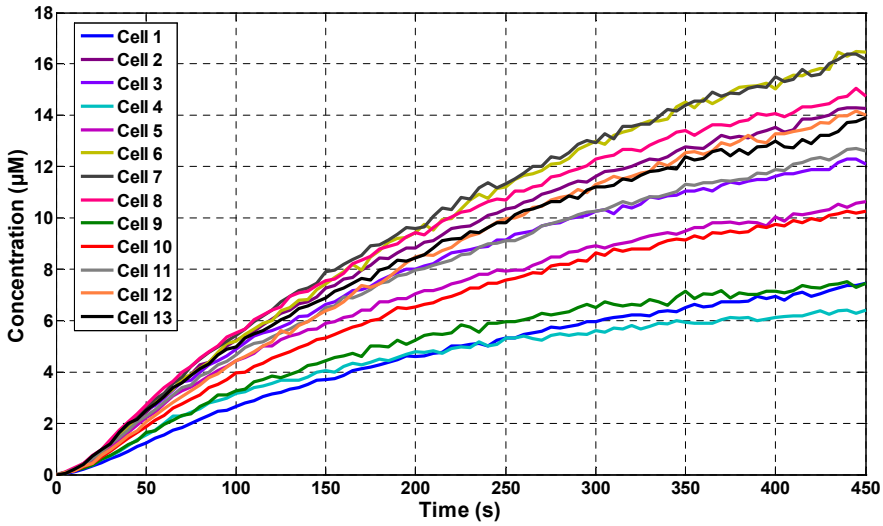


Figure 4.11: Observations from the nucleus (L_n) for the 13 individual cells that represent the full range of heterogeneity observed in response to the drug.

Table 4.3 represents *normalised* AUC(L_n, D), D = 10 µM, values (with respect to the maximum AUC(L_n, 10 µM)) that correspond to the 13 individual cells. Although Cell 4 has the lowest normalised AUC(L_n, D) value, Cell 1 has been selected as a low loading cell for parameter estimation due to noise in the measurements in the cytoplasm (i.e., $L_c + H_c$)

for Cell 4. Cell 7 has been chosen as a high loading cell as the concentration of TPT_L in the nucleus is relatively high.

Table 4.4: Best parameter estimates for averaged data obtained for the model Equations (4.12), estimated using TPLSM data. The cellular hydrolysis rate constants are fixed at values obtained from Table 4.2 (see Subsection 4.2.2) for pH = 7.2 buffered solution where: $k_{om} = k_{oc} = 1.6 \times 10^{-4}$ s^{-1} and $k_{cm} = k_{cc} = 3 \times 10^{-4}$ s^{-1}.

Parameter	Value	SDLN	5%	95%
k_i (s^{-1})	2.3599×10^{-2}	0.0531	2.1625×10^{-2}	2.5754×10^{-2}
k_e (s^{-1})	1.0266×10^{-2}	0.0594	9.3106×10^{-3}	1.1321×10^{-2}
k_b (s^{-1} μM^{-1})	1.9483×10^{-4}	0.8778	4.5975×10^{-5}	8.2564×10^{-4}
k_{mi} (s^{-1})	1.4539×10^{-6}	0.0374	1.3671×10^{-6}	1.5463×10^{-6}
k_{mo} (s^{-1})	9.1051×10^{-2}	0.0277	8.7001×10^{-2}	9.5290×10^{-2}
k_{dl} (s^{-1})	3.5101×10^{-2}	0.1821	2.6015×10^{-2}	4.7360×10^{-2}
k_{dh} (s^{-1})	2.0000×10^{-4}	1.2927	2.3852×10^{-5}	1.6770×10^{-3}
v_0	1.4778×10^{-5}	0.0539	1.3525×10^{-5}	1.6147×10^{-5}
B_T (μM)	$1.3412\times10^{+2}$	0.952	$2.8015\times10^{+1}$	$6.4209\times10^{+2}$
V_{m1} (s^{-1} μM)	1.6071×10^{-2}	0.1746	1.2060×10^{-2}	2.1417×10^{-2}
V_{m2} (s^{-1} μM)	4.0086×10^{-3}	0.1465	3.1500×10^{-3}	5.1013×10^{-3}
k_{m1} (μM)	1.1400×10^{-6}	0.3981	5.9221×10^{-7}	2.1945×10^{-6}
k_{m2} (μM)	9.6828×10^{-6}	0.1616	7.4231×10^{-6}	1.2630×10^{-5}
RSS = $2.3423\times10^{+1}$				

The fitted (estimated) parameter values, with estimates for their confidence levels are presented in Tables 4.4-4.6. These parameter estimates correspond to the averaged data fitting, high loader fitting and low loader fitting, respectively. In Table 4.4, which corresponds to the averaged data fitting, it is seen that all parameters are well-determined by the microscopy data. In addition, ten parameters, namely, k_i, k_e, k_{mi}, k_{mo}, k_{dl}, v_0, V_{m1}, V_{m2}, k_{m1} and k_{m2}, have low SDLN values corresponding to high confidence in their values. However, the estimated parameter values for k_b, k_{dh} and B_T have high SDLN values due to the low signal-to-noise ratio in the experimental data. The value of $CORI_{mean} = 2.00848$ (in the averaged fitting, Table 4.4) indicates that the residuals are acceptable and evenly distributed above and below the simulated output of the model and therefore, the resulting fit is considered to be good.

Table 4.5: Best parameter estimates for high loading cell data obtained for the model Equations (4.12), estimated using TPLSM data. The cellular hydrolysis rate constants are fixed at values obtained from Table 4.2 (see Subsection 4.2.2) for pH = 7.2 buffered solution where: $k_{om} = k_{oc} = 1.6 \times 10^{-4}$ s^{-1} and $k_{cm} = k_{cc} = 3 \times 10^{-4}$ s^{-1}.

Parameter	Value	SDLN	5%	95%
k_i (s^{-1})	1.7762×10^{-1}	0.1390	1.4130×10^{-1}	2.2327×10^{-1}
k_e (s^{-1})	9.8046×10^{-3}	0.3720	5.3167×10^{-3}	1.8081×10^{-2}
k_b (s^{-1} μM^{-1})	5.3077×10^{-4}	0.1611	4.0722×10^{-4}	6.9180×10^{-4}
k_{mi} (s^{-1})	1.3585×10^{-6}	0.0422	1.2674×10^{-6}	1.4561×10^{-6}
k_{mo} (s^{-1})	1.0108×10^{-1}	0.0521	9.2774×10^{-2}	1.1013×10^{-1}
k_{dl} (s^{-1})	2.8264×10^{-2}	0.1322	2.2741×10^{-2}	3.5127×10^{-2}
k_{dh} (s^{-1})	1.1245×10^{-5}	NWD	-	-
k_{m2} (μM)	5.9031×10^{-2}	0.5286	2.4741×10^{-2}	1.4085×10^{-1}
v_0	1.1951×10^{-5}	0.0866	1.0364×10^{-5}	1.3782×10^{-5}
B_T (μM)	$4.7215\times10^{+1}$	0.3006	$2.8795\times10^{+1}$	$7.7418\times10^{+1}$
V_{m1} (s^{-1} μM)	$1.4157\times10^{+1}$	0.1075	$1.1862\times10^{+1}$	$1.6895\times10^{+1}$
V_{m2} (s^{-1} μM)	1.4245×10^{-1}	0.1187	1.1717×10^{-1}	1.7317×10^{-1}
k_{m1} (μM)	3.6341	0.2181	2.5387	5.2021

RSS = $3.2520\times10^{+1}$

Table 4.6: Best parameter estimates for low loading cell data obtained for the model Equations (4.12), estimated using TPLSM data. The cellular hydrolysis rate constants are fixed at values obtained from Table 4.2 (see Subsection 4.2.2) for pH = 7.2 buffered solution where: $k_{om} = k_{oc} = 1.6 \times 10^{-4}$ s^{-1} and $k_{cm} = k_{cc} = 3 \times 10^{-4}$ s^{-1}.

Parameter	Value	SDLN	5%	95%
k_i (s^{-1})	7.3625×10^{-2}	0.1893	5.3925×10^{-2}	1.0052×10^{-1}
k_e (s^{-1})	7.8178×10^{-3}	0.0478	7.2268×10^{-3}	8.4571×10^{-3}
k_b (s^{-1} μM^{-1})	1.7577×10^{-4}	0.2127	1.2387×10^{-4}	2.4941×10^{-4}
k_{mi} (s^{-1})	9.1689×10^{-7}	0.0844	7.9808×10^{-7}	1.0534×10^{-6}
k_{mo} (s^{-1})	9.0915×10^{-2}	0.0418	8.4878×10^{-2}	9.7381×10^{-2}
k_{dl} (s^{-1})	4.0550×10^{-2}	0.0832	3.5361×10^{-2}	4.6501×10^{-2}
k_{dh} (s^{-1})	1.1362×10^{-5}	NWD		
k_{m2} (μM)	7.5653×10^{-2}	0.2066	5.3851×10^{-2}	1.0628×10^{-1}
v_0	9.3481×10^{-6}	0.1096	7.8054×10^{-6}	1.1196×10^{-5}
B_T (μM)	$1.2689\times10^{+2}$	0.2292	$8.7040\times10^{+1}$	$1.8500\times10^{+2}$
V_{m1} (s^{-1} μM)	7.3930	0.1341	5.9293	9.2181
V_{m2} (s^{-1} μM)	6.1983×10^{-2}	0.0615	5.6019×10^{-2}	6.8583×10^{-2}
k_{m1} (μM)	8.4899×10^{-3}	0.1856	6.2558×10^{-3}	1.1522×10^{-2}

RSS = $3.8679\times10^{+1}$

In Tables 4.5 and 4.6, which correspond to the high loader fitting and low loader fitting, respectively, it is seen that the same one of the thirteen parameters (k_{dh}) is not well-determined (NWD) by the data. That is, small deviations from the obtained estimates for the parameter k_{dh} result in similar fits. The final estimates for the unknown model parameter k_{dh} were gained from multiple fits with a range of initial guesses. All well determined parameters (in Tables 4.5 and 4.6) have low SDLN values corresponding to high confidence in their estimated values. The one exception is k_{m2} in the high loader fitting and is due to low signal-to-noise ratio in the experimental data. The values of $CORI_{mean} = 0.75542$ (in the high loader fitting) and $CORI_{mean} = 1.61803$ (in the low loader fitting) indicate that the residuals (in both fits) are acceptable and evenly distributed above and below the simulated output of the model and therefore, the resulting fits are good.

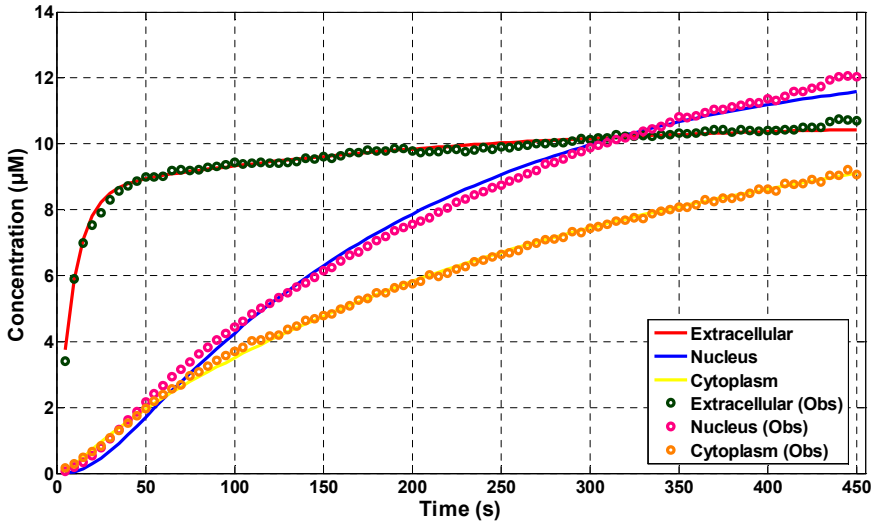


Figure 4.12: Simulated output for averaged data for the model Equations (4.12), with parameters taking values in Table 4.4, plotted (solid) against experimental data (circles).

Based on Tables 4.4 and 4.5, the rate constants for the flow of both $L_e(t)$ and $L_m(t)$ into the cytoplasm and extracellular pool (respectively) are higher in the high loader than the low loader, suggesting that these parameters play important roles in transporting the active form of the drug (TPT_L) to the cellular pool. In addition, the binding constant (k_b) at which TPT_L binds to DNA increases for the high loader, which offers another reason why the cells receive more pharmacologically active drug. The binding affinity of the drug (k_{m1}) to the efflux pump is around 428 times higher in the low loading cell, therefore, the drug transporter mainly resists the anti-cancer agent TPT. The rate constant for dissociation of bound drug (k_{dl}) is relatively low for the high loading cell, providing another reason for the

retention of the active form of the drug bound to DNA for these cells. Although the concentration of TPT_L in the nucleus for the high loader is relatively high compared to that for the low loader, the value of B_T in the low loading cell is larger. This suggests that the efflux pumping mechanism plays a key role in the resistance of anti-cancer drugs.

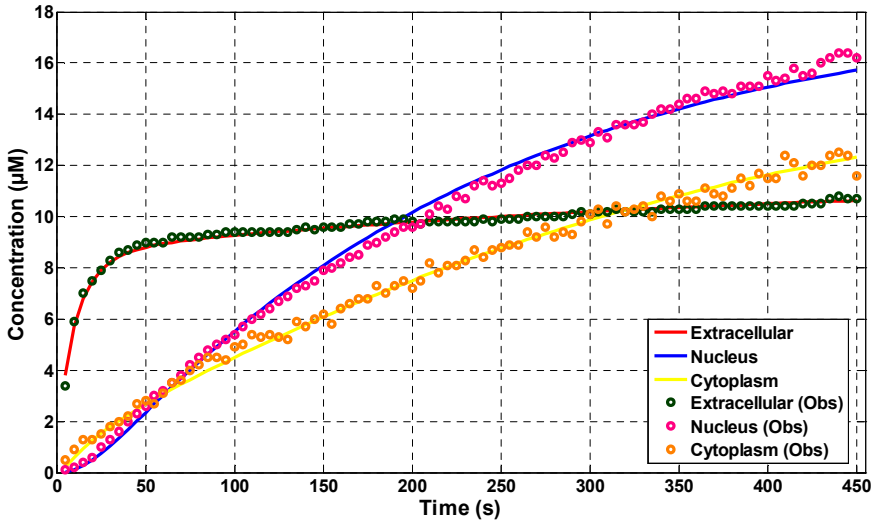


Figure 4.13: Simulated output for high loading cell data for the model Equations (4.12), with parameters taking values in Table 4.5, plotted (solid) against experimental data (circles).

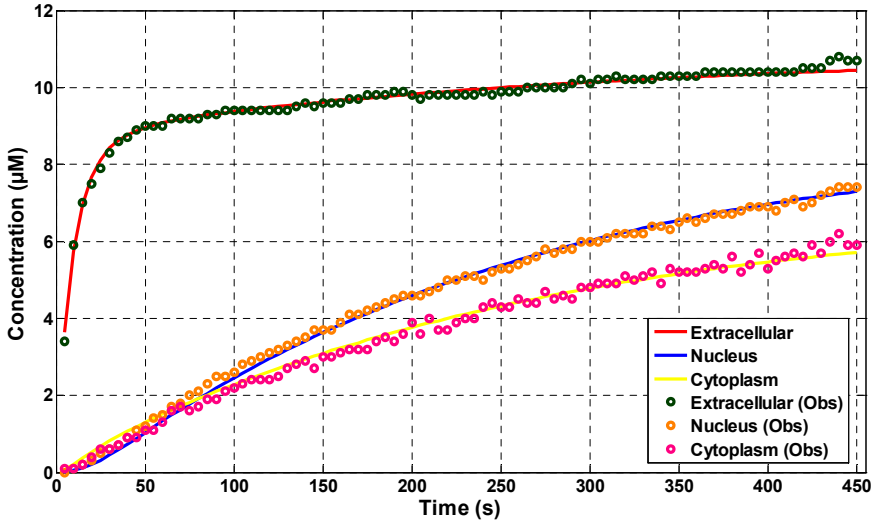


Figure 4.14: Simulated output for low loading cell data for the model Equations (4.12), with parameters taking values in Table 4.6, plotted (solid) against experimental data (circles).

The plots shown in Figures 4.12-4.14 represent the model fits for the averaged data, high loader and low loader, respectively. These fits show that there is close reproduction of the

data by the simulated output from the model with parameters taking values from Tables 4.4-4.6. Figures 4.15-4.17 show normal probability plots of the three fittings, i.e. average, high loader and low loader, and this is used to test whether or not the weighted residuals are normally distributed. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale. The figures are consistent with the assumption of normally distributed errors. This is an indicator for the appropriateness of the standard error σ_i used for each observation y_i.

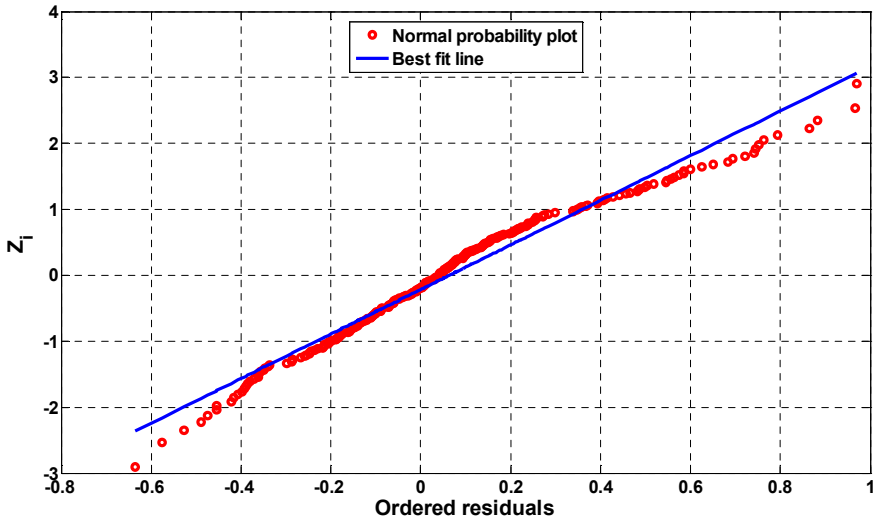


Figure 4.15: A normal probability plot for the residuals in the weighted non-linear least-squares fit (averaged data) of the extended model Equations (4.12) to TPLSM data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

The estimated correlation matrices for the well-determined parameters in each of the fits are given in Tables 4.7-4.9 (average data, high loader and low loader fittings, respectively). The parameter pairings that exhibit noticeable (highest) correlation are: k_i and k_{mi} (-0.935), k_i and v_0 (-0.938), k_{mi} and v_0 (0.933), k_i and k_{mo} (0.577), k_b and B_T (-0.995), k_b and k_{dl} (-0.707), B_T and k_{dl} (0.774), B_T and k_{dh} (0.716), V_{m1} and B_T (-0.644) and V_{m1} and V_{m2} (0.631). The pairs that involve B_T and/or k_b relate to the association/dissociation of TPT_L to/from DNA target. The pairs that involve V_{m1} and/or V_{m2} relate to the drug resistance mechanisms in the cell, namely, ALDH and the BCRP/ABCG2 transporter. The rest of the pairings relate to the crossing of the medium/extracellular region or extracellular region/cytoplasm borders in which the modelling of this involves the ratio v_0. Therefore, this parameter, i.e., v_0, is important and other parameters depend on it in the model. The high correlation of the

logarithms of these parameters indicates the significance of their product or ratio (i.e., drug affinity) for the fitting procedure.

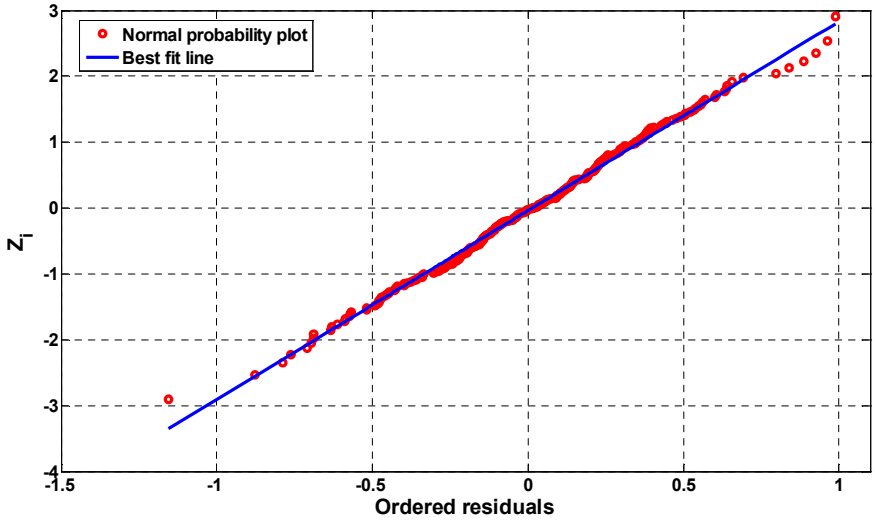


Figure 4.16: A normal probability plot for the residuals in the weighted non-linear least-squares fit (high loading cell data) of the extended model Equations (4.12) to TPLSM data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

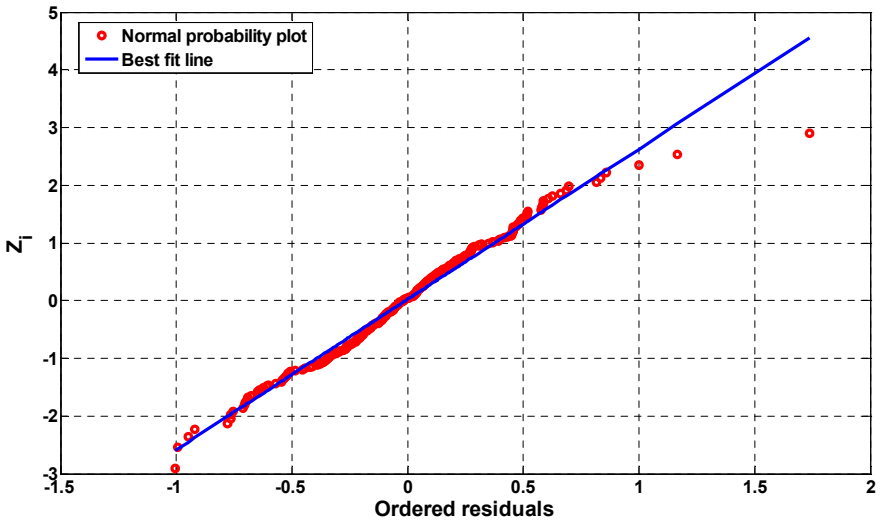


Figure 4.17: A normal probability plot for the residuals in the weighted non-linear least-squares fit (low loading cell data) of the extended model Equations (4.12) to TPLSM data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

Table 4.7: The estimated correlation matrix for the well-determined parameters for averaged data fitting

	k_i	k_e	k_b	k_{mi}	k_{mo}	k_{dl}	k_{dh}	v_0	B_T	V_{m1}	V_{m2}	k_{m1}	k_{m2}
k_i	1	0.317	-0.204	-0.838	0.577	0.26	0.491	-0.926	0.233	0.101	-0.234	-0.201	-0.315
k_e	0.317	1	-0.529	0	-0.098	-0.088	0.73	0.05	0.47	0.43	0.478	-0.471	0.184
k_b	-0.204	-0.529	1	0.038	0.023	-0.707	-0.74	0.021	-0.995	0.045	0.183	-0.152	0.104
k_{mi}	-0.838	0	0.038	1	-0.22	-0.301	-0.265	0.871	-0.087	0.008	0.401	0.073	0.306
k_{mo}	0.577	-0.098	0.023	-0.22	1	0.158	0.105	-0.67	0.008	-0.149	-0.226	0.011	-0.305
k_{dl}	0.26	-0.088	-0.707	-0.301	0.158	1	0.314	-0.311	0.774	-0.339	-0.562	0.372	-0.251
k_{dh}	0.491	0.73	-0.74	-0.265	0.105	0.314	1	-0.258	0.716	0.017	-0.177	-0.187	-0.251
v_0	-0.926	0.05	0.021	0.871	-0.67	-0.311	-0.258	1	-0.074	0.084	0.421	0.049	0.387
B_T	0.233	0.47	-0.995	-0.087	0.008	0.774	0.716	-0.074	1	-0.086	-0.242	0.178	-0.127
V_{m1}	0.101	0.43	0.045	0.008	-0.149	-0.339	0.017	0.084	-0.086	1	0.631	-0.189	0.292
V_{m2}	-0.234	0.478	0.183	0.401	-0.226	-0.562	-0.177	0.421	-0.242	0.631	1	-0.487	0.638
k_{m1}	-0.201	-0.471	-0.152	0.073	0.011	0.372	-0.187	0.049	0.178	-0.189	-0.487	1	-0.4
k_{m2}	-0.315	0.184	0.104	0.306	-0.305	-0.251	-0.251	0.387	-0.127	0.292	0.638	-0.4	1

Table 4.8: The estimated correlation matrix for the well-determined parameters for high loading cell data fitting

	k_i	k_e	k_b	k_{mi}	k_{mo}	k_{dl}	k_{m2}	V_0	B_T	V_{m1}	V_{m2}	k_{m1}
k_i	1	0.45	-0.481	0.014	0.341	-0.324	-0.043	-0.176	0.259	0.424	0.676	-0.754
k_e	0.45	1	-0.794	0.136	-0.138	0.033	-0.693	0.164	0.895	-0.241	0.678	-0.195
k_b	-0.481	-0.794	1	-0.029	-0.141	-0.104	0.246	0.062	-0.892	0.217	-0.56	0.281
k_{mi}	0.014	0.136	-0.029	1	-0.526	0.113	-0.352	0.869	0.089	0.655	0.642	-0.288
k_{mo}	0.341	-0.138	-0.141	-0.526	1	-0.317	0.529	-0.877	-0.113	-0.228	-0.28	-0.146
k_{dl}	-0.324	0.033	-0.104	0.113	-0.317	1	-0.165	0.243	0.362	-0.165	-0.094	-0.21
k_{m2}	-0.043	-0.693	0.246	-0.352	0.529	-0.165	1	-0.509	-0.527	0.109	-0.459	-0.002
v_0	-0.176	0.164	0.062	0.869	-0.877	0.243	-0.509	1	0.117	0.509	0.534	-0.09
B_T	0.259	0.895	-0.892	0.089	-0.113	0.362	-0.527	0.117	1	-0.373	0.521	-0.17
V_{m1}	0.424	-0.241	0.217	0.655	-0.228	-0.165	0.109	0.509	-0.373	1	0.534	-0.606
V_{m2}	0.676	0.678	-0.56	0.642	-0.28	-0.094	-0.459	0.534	0.521	0.534	1	-0.591
k_{m1}	-0.754	-0.195	0.281	-0.288	-0.146	-0.21	-0.002	-0.09	-0.17	-0.606	-0.591	1

Table 4.9: The estimated correlation matrix for the well-determined parameters for high loading cell data fitting

	k_i	k_e	k_b	k_{mi}	k_{mo}	k_{dl}	k_{m2}	v_0	B_T	V_{m1}	V_{m2}	k_{m1}
k_i	1	0.236	0.31	-0.935	0.389	-0.325	-0.365	-0.938	-0.367	0.521	0.65	0.158
k_e	0.236	1	-0.431	-0.151	0.033	0.089	-0.542	-0.137	0.411	-0.176	0.126	0.181
k_b	0.31	-0.431	1	-0.43	0.058	-0.201	0.383	-0.376	-0.948	0.599	0.131	0.176
k_{mi}	-0.935	-0.151	-0.43	1	-0.221	0.276	0.346	0.933	0.456	-0.559	-0.423	-0.331
k_{mo}	0.389	0.033	0.058	-0.221	1	-0.181	0.089	-0.556	-0.109	0.115	0.072	-0.185
k_{dl}	-0.325	0.089	-0.201	0.276	-0.181	1	-0.249	0.297	0.501	-0.367	-0.33	0.472
k_{m2}	-0.365	-0.542	0.383	0.346	0.089	-0.249	1	0.267	-0.428	-0.146	-0.158	-0.58
v_0	-0.938	-0.137	-0.376	0.933	-0.556	0.297	0.267	1	0.416	-0.51	-0.379	-0.218
B_T	-0.367	0.411	-0.948	0.456	-0.109	0.501	-0.428	0.416	1	-0.644	-0.218	0.005
V_{m1}	0.521	-0.176	0.599	-0.559	0.115	-0.367	-0.146	-0.51	-0.644	1	0.341	-0.019
V_{m2}	0.65	0.126	0.131	-0.423	0.072	-0.33	-0.158	-0.379	-0.218	0.341	1	-0.276
k_{m1}	0.158	0.181	0.176	-0.331	-0.185	0.472	-0.58	-0.218	0.005	-0.019	-0.276	1

In order to study the variations across a high loader and a low loader, the extended drug kinetics model (Equation (4.12)) was fitted to the high loading cell data and low loading cell data simultaneously as shown in Figures 4.18-4.20. This fitting was performed by assuming that such variations across cells are only due to differences between the parameters relating to the ALDH enzyme and the BCRP/ABCG2 drug transporter. That is, the parameters, k_i, k_e, k_b, k_{mi}, k_{mo}, k_{dl}, k_{dh}, v_0 and B_T, are fixed in both cells and the parameters V_{m1}, V_{m2}, k_{m1} and k_{m2} vary between both cells.

Table 4.10: Best parameter estimates for simultaneous fitting between high loading cell data and low loading cell obtained for the model Equations (4.12), estimated using TPLSM data. The cellular hydrolysis rate constants are fixed at values obtained from Table 4.2 (see Subsection 4.2.2) for pH = 7.2 buffered solution where: $k_{om} = k_{oc} = 1.6 \times 10^{-4}$ s^{-1} and $k_{cm} = k_{cc} = 3 \times 10^{-4}$ s^{-1}.

Common parameters

Parameter	Value	SDLN	5%	95%
k_i (s^{-1})	1.04×10^{-1}	0.152	8.13×10^{-2}	1.34×10^{-1}
k_e (s^{-1})	8.86×10^{-1}	0.2314	6.06×10^{-1}	1.30
k_b (s^{-1} μM^{-1})	1.34×10^{-3}	0.5855	5.13×10^{-4}	3.52×10^{-3}
k_{mi} (s^{-1})	2.83×10^{-6}	0.0498	2.61×10^{-6}	3.07×10^{-6}
k_{mo} (s^{-1})	7.97×10^{-2}	0.0303	7.58×10^{-2}	8.38×10^{-2}
k_{dl} (s^{-1})	3.57×10^{-3}	0.0745	3.16×10^{-3}	4.03×10^{-3}
k_{dh} (s^{-1})	1.8331×10^{-7}	NWD		
v_0	3.43×10^{-5}	0.0679	3.07×10^{-5}	3.84×10^{-5}
B_T (μM)	$9.28\times10^{+1}$	0.4708	$4.28\times10^{+1}$	$2.01\times10^{+2}$

Parameters relating to the high loading cell

Parameter	Value	SDLN	5%	95%
V_{m1} (s^{-1} μM)	$2.47\times10^{+1}$	0.0937	$2.12\times10^{+1}$	$2.88\times10^{+1}$
V_{m2} (s^{-1} μM)	2.70×10^{-1}	0.0859	2.35×10^{-1}	3.11×10^{-1}
k_{m1} (μM)	$2.74\times10^{+1}$	0.1378	$2.18\times10^{+1}$	$3.43\times10^{+1}$
k_{m2} (μM)	9.20×10^{-4}	0.2274	6.33×10^{-4}	1.34×10^{-3}

Parameters relating to the low loading cell

Parameter	Value	SDLN	5%	95%
V_{m1} (s^{-1} μM)	$3.13\times10^{+1}$	0.0767	$2.76\times10^{+1}$	$3.55\times10^{+1}$
V_{m2} (s^{-1} μM)	5.78×10^{-1}	0.0496	5.33×10^{-1}	6.27×10^{-1}
k_{m1} (μM)	3.16	0.277	2.00	4.98
k_{m2} (μM)	8.39×10^{-2}	0.2263	5.79×10^{-2}	1.22×10^{-1}

RSS = $1.46139\times10^{+2}$

The fitted (estimated) parameter values with estimates for their confidences are presented in Table 4.10, where it is seen that the parameter k_{dh} is NWD by the data. That is, small deviations from the obtained estimate for the unknown model parameter k_{dh} will result in similar fits. The estimated parameter values for k_b and B_T have high SDLN values due to the low signal-to-noise ratio in the experimental data from the nucleus and cytoplasm, particularly for the high loader (see Figures 4.19 and 4.20). However, the rest of the parameter estimates in Table 4.10 have low SDLN values corresponding to high confidence in their values. The value of $\text{CORI}_{\text{mean}} = 4.29423$ (in the simultaneous fitting) indicates that the residuals are acceptable and evenly distributed above and below the simulated output of the model and therefore, the resulting fit is good.

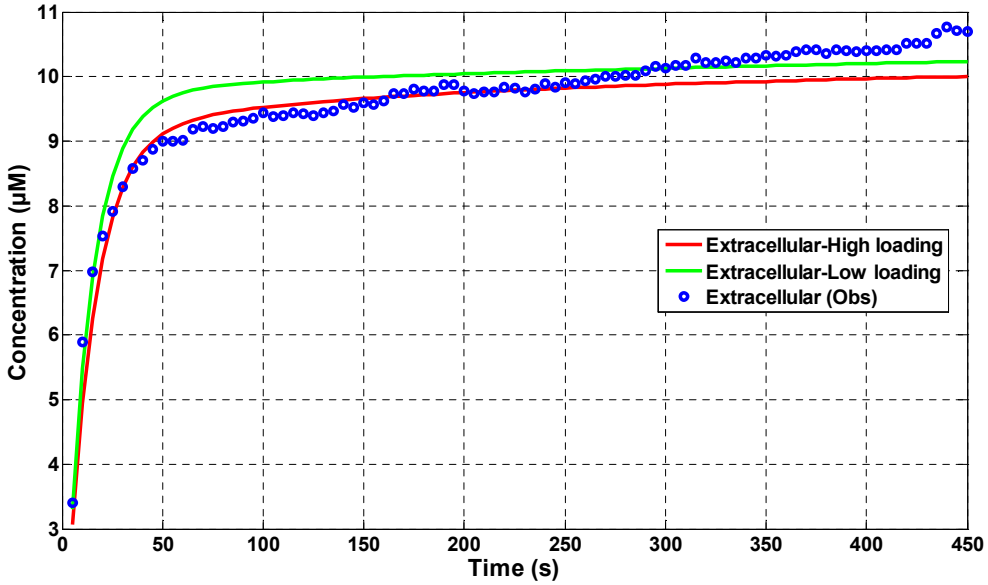


Figure 4.18: Simulated output of the extracellular region for simultaneous fitting of the high loading cell data and the low loading cell data for the model Equations (4.12), with parameters taking values in Table 4.10, plotted (solid) against experimental data (circles).

Based on Table 4.10, the binding affinity of the inactive form of the drug (k_{m1}) to the efflux pump is higher in the low loading cell, as expected, the efflux pumping mechanism plays a key role in the resistance of anti-cancer drugs. In addition, the results obtained in the simultaneous fitting are consistent with the previous results, that is, the drug transporter (BCRP/ABCG2) is a key factor in resisting TPT. The plots shown in Figures 4.18-4.20 represent the simultaneous model fits for the extracellular region, nucleus and cytoplasm, respectively. These fits show that there is close reproduction of the data by the simulated output from the model with parameters taking values from Table 4.10. Figure 4.21 shows a

normal probability plot of the simultaneous fitting. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale. The figures are consistent with the assumption of normally distributed errors. This is an indicator for the appropriateness of the standard error σ_i used for each observation y_i.

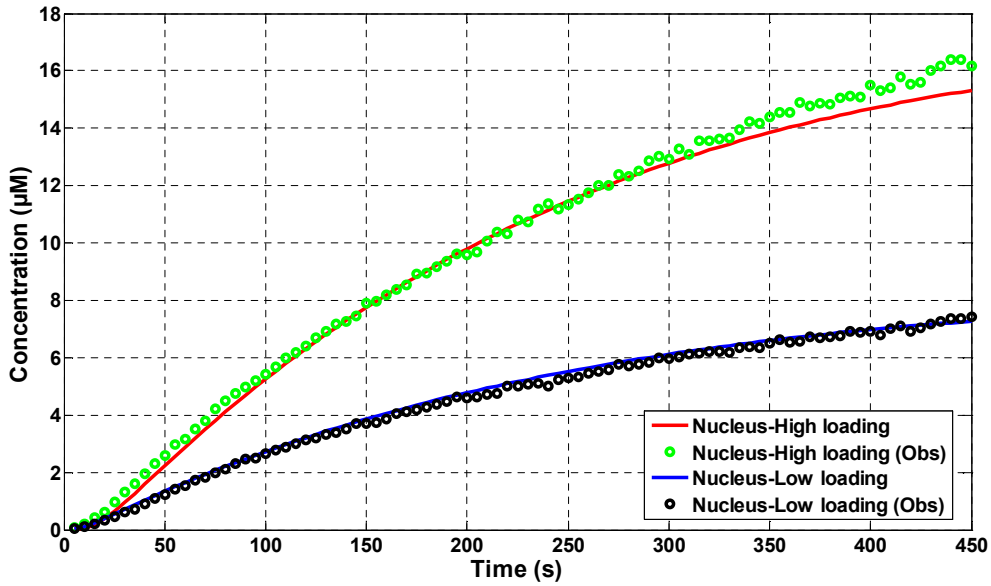


Figure 4.19: Simulated output of the nucleus for simultaneous fitting of the high loading cell data and the low loading cell data for the model Equations (4.12), with parameters taking values in Table 4.10, plotted (solid) against experimental data (circles).

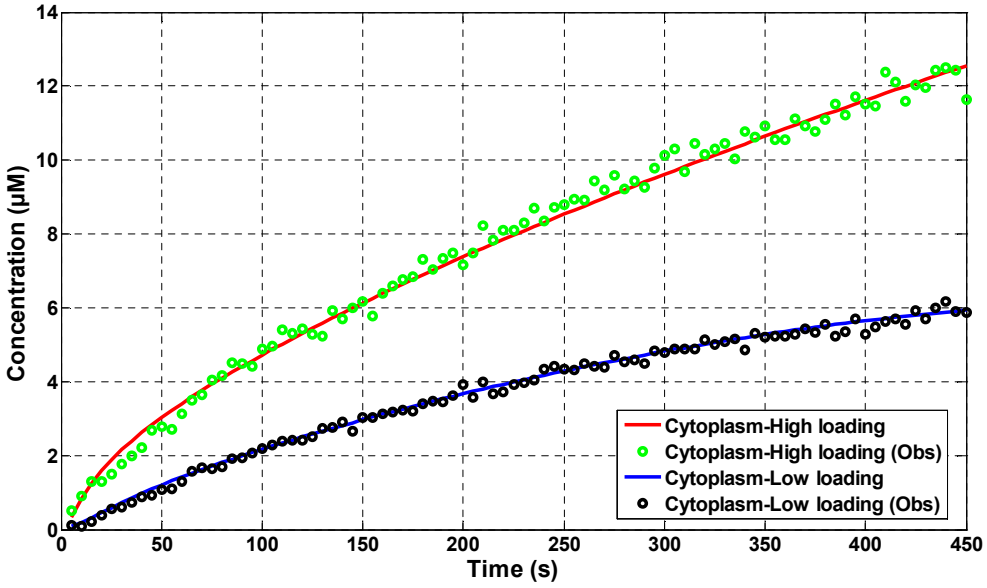


Figure 4.20: Simulated output of the cytoplasm for simultaneous fitting of the high loading cell data and the low loading cell data for the model Equations (4.12), with parameters taking values in Table 4.10, plotted (solid) against experimental data (circles).

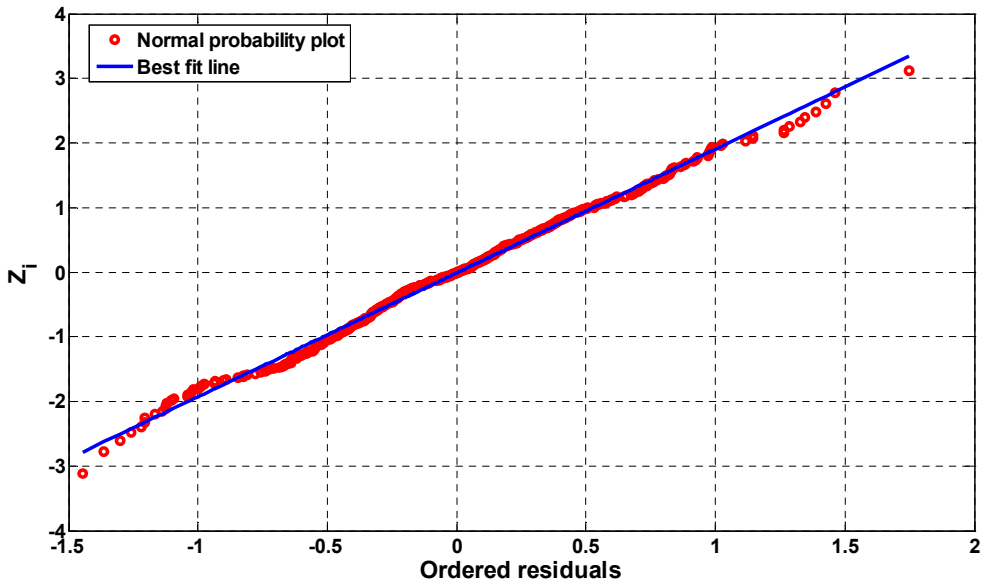


Figure 4.21: A normal probability plot for the residuals in the weighted non-linear least-squares fit (low loading cell and high loading cell data simultaneously) of the extended model Equations (4.12) to TPLSM data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

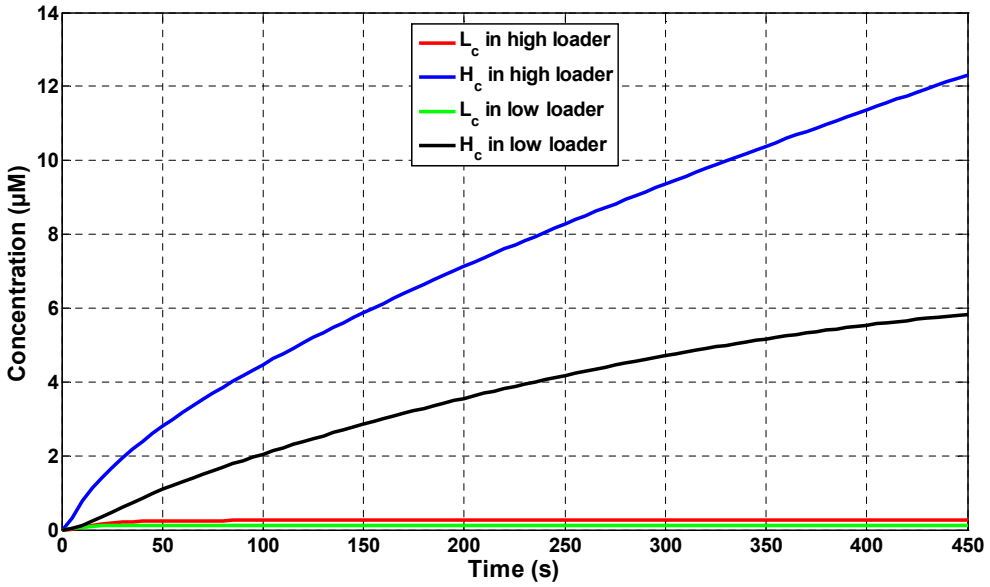


Figure 4.22: Simulated output for L_c and H_c of a high loader and a low loader for the model Equations (4.12), with parameters taking values in Table 4.10.

An advantage of mathematical modelling is to infer from the model information relating to state variables of the system (in this case, the extended cell cycle model) that are not directly measured (i.e., parts of the system that are not accessible) [13]. For example, in the cytoplasm it is the total concentration of the drug ($L_c + H_c$) that is directly observed (using

microscopy, see Section 4.5.1). However, using the model, it is possible to separately predict the concentrations of TPT_L and TPT_H in the cytoplasm (i.e., L_c and H_c, respectively). Figure 4.22 shows simulated output from the extended drug kinetics model (with parameters taking values in Table 4.10) for L_c and H_c in a high loader and a low loader. The concentrations of the active form of the drug in the cytoplasm (L_c) are almost indistinguishable between the high loader and low loader and converge over time. In addition, the concentrations of the inactive form of the drug in the cytoplasm (H_c) differ noticeably between both cells and diverge over time. As expected, the efflux pumping mechanism resulting from the BCRP/ABCG2 transporter accounts for reducing the concentration of TPT in the cell.

4.6 Sensitivity analysis

It was shown from the previous section that the main differences between the model predictions, for the high loader and low loader, using the parameter values in Table 4.10 is the concentration of TPT_H in the cytoplasm (i.e., H_c) and the concentration of TPT_L in the nucleus (L_n). Other predicted output does not seem to be sensitive to the different values assumed for the binding affinity (k_{m1}) of TPT_H (in the cytoplasm) to the BCRP/ABCG2 transporter and the binding affinity (k_{m2}) of TPT_L (in the cytoplasm) to the enzyme ALDH. In this section the impact of varying the values for k_{m1} and k_{m2} is considered with respect to individual state variable and predicted dose response at the target (nuclear DNA).

4.6.1 Sensitivity analysis of the single-cell model

The impact of varying the values for the binding affinity of TPT_H to the transporter, k_{m1}, is shown in Figure 4.23 Parameters in the single-cell model were assigned values from the high loader fit (for demonstration purposes) in Table 4.10, and the parameter k_{m1} was varied in the range of 2.74×10^{-1}-$2.74\times10^{+2}$ μM. The values are: (a) 2.74×10^{-1} μM, (b) 2.74 μM, (c) $2.74\times10^{+1}$ μM and (d) $2.74\times10^{+2}$ μM. The only model variable exhibiting noticeable variation in simulation was H_c when varying k_{m1}. Decreasing the binding affinity of TPT_H to the BCRP/ABCG2 transporter (i.e., increasing the value of the

parameter k_{m1}) results in an increase in the concentration of TPT_H in the cytoplasm. This is not surprising since the drug is retained in the cytoplasm and is not flushed out of the cell when the binding affinity decreases.

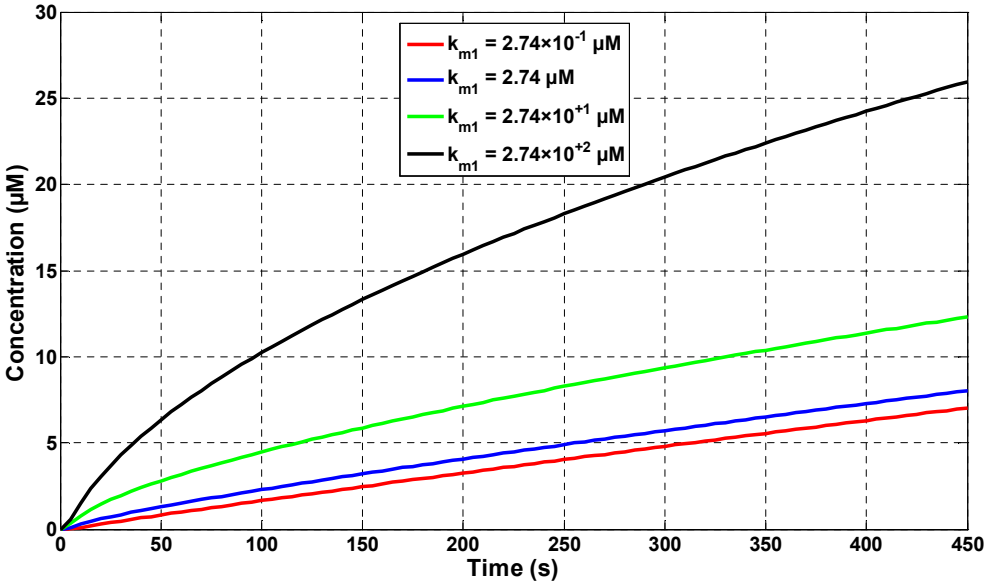


Figure 4.23: Simulated output for H_c from the model Equations (4.12), with parameters taking values in Table 4.10, while varying the binding affinity of TPT_H to the BCRP/ABCG2 transporter k_{m1} in the ranges (a) 2.74×10^{-1} µM to (d) $2.74\times10^{+2}$ µM.

The impact of varying the values for the binding affinity of TPT_L (in the cytoplasm) to the enzyme ALDH, k_{m2}, is shown in Figures 4.24 and 4.25. Parameters in the single-cell model were assigned values from the high loader fit in Table 4.10, and the parameter k_{m2} was varied in the range of 9.2×10^{-4}-$9.2\times10^{+2}$ µM. The values are: (a) 9.2×10^{-4} µM, (b) 9.2×10^{-2} µM, (c) 9.2 µM and (d) $9.2\times10^{+2}$ µM. The only model variables exhibiting noticeable variation in simulation were H_c (Figure 4.24) and L_n (Figure 4.25) when varying k_{m2}. As expected, decreasing the binding affinity of TPT_L (in cytoplasm) to the enzyme ALDH (i.e., increasing the value of the parameter k_{m2}) results in a decrease in the concentration of TPT_H in the cytoplasm (H_c), as shown in Figure 4.24. That is, decreasing the binding affinity results in a decrease in the drug resistance caused by the enzyme ALDH. In addition, increasing the value of the parameter k_{m2} (i.e., decreasing the binding affinity of L_c to the enzyme) results in an increase in the drug bound to nuclear DNA (Figure 4.25). The lack of sensitivity of the state variable L_c to variation in k_{m2} suggests that this is not the key process for the presence of TPT_L in the cytoplasm. However, the passive influx/efflux

of TPT_L across the plasma membrane in addition to its association/dissociation are more crucial. These results confirm that the enzyme ALDH is another key factor in resisting TPT.

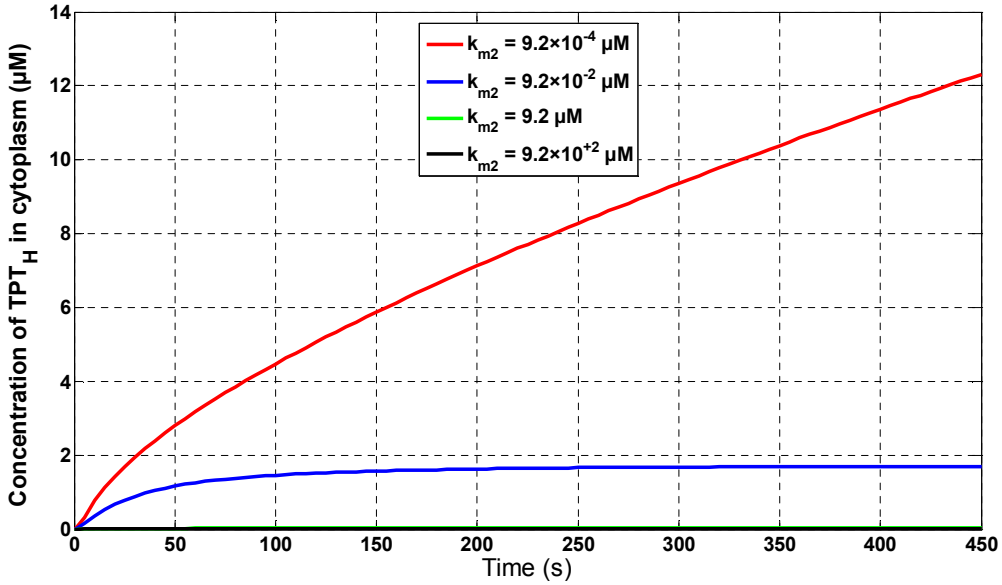


Figure 4.24: Simulated output for H_c from the model Equations (4.12), with parameters taking values in Table 4.10, while varying the binding affinity of TPT_L to the enzyme ALDH k_{m2} in the ranges (a) 9.2×10^{-4} μM to (d) $9.2\times10^{+2}$ μM.

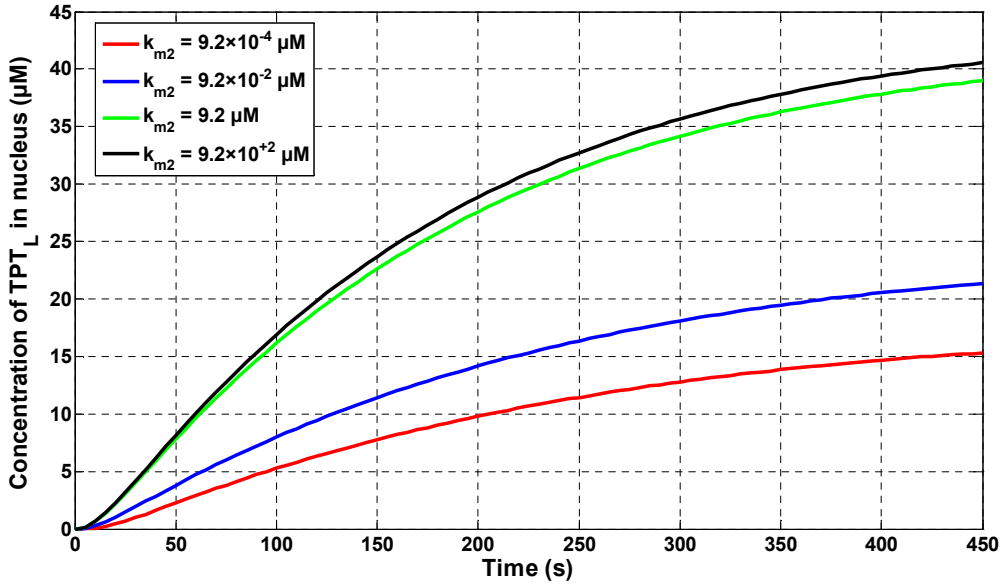


Figure 4.25: Simulated output for L_n from the model Equations (4.12), with parameters taking values in Table 4.10, while varying the binding affinity of TPT_L to the enzyme ALDH k_{m2} in the ranges (a) 9.2×10^{-4} μM to (d) $9.2\times10^{+2}$ μM.

4.6.2 Sensitivity of dose response

As explained in Section 4.5, the impact of the dose (following the drug administration) can be measured by calculating the AUC for L_n corresponding to the initial dose D, i.e., AUC(L_n, D). Similar to the approach used in Evans *et al.* [13], the therapeutic impact of a dose was measured by AUC(L_n, D) over the first hour (3600 s) following drug administration. Simulations were performed with D taking values from 0.1 μM to 20 μM for the high loading cell (Figure 4.26) and low loading cell (Figure 4.27).

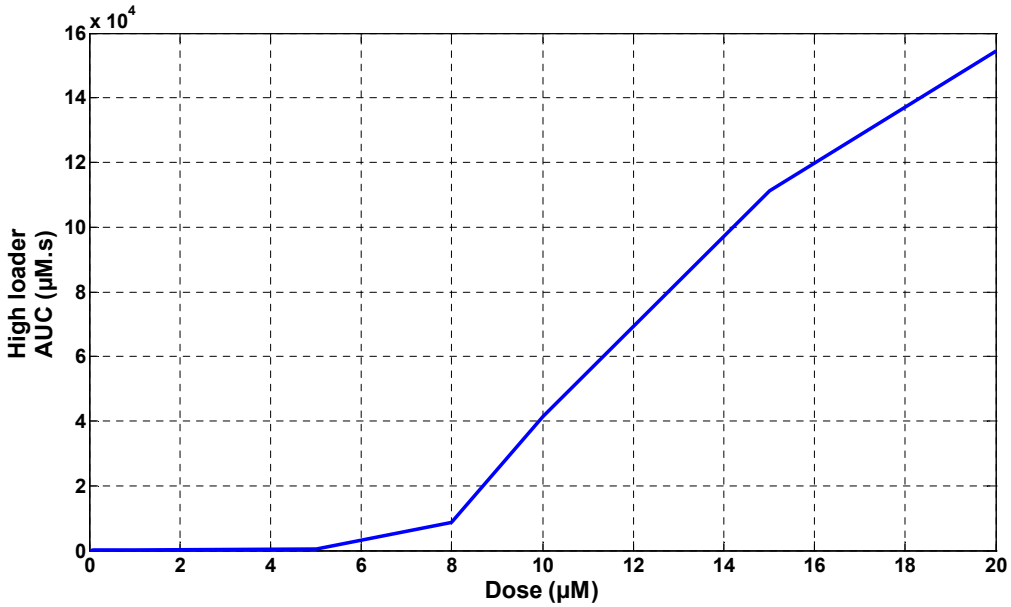


Figure 4.26: AUC for L_n over 3600 seconds (1 hour) against dose D using the parameter values for the high loader.

For both the high loading cell and low loading cell, at low values for the dose D the impact is relatively low. As the value of the dose D increases up to 8 μM (in the high loading cell) and 10 μM (in the low loading cell) the rate of change of AUC with respect to dose starts increasing. In the low loading cell (Figure 4.27) increasing the dose D results in an increase in the rate of change of AUC with respect to dose, however, in the high loading cell at around D = 15 μM, the slope of the impact starts decreasing. This suggests that there is a maximum damaging effect on cells and further increase in the treatment regimen will not have further therapeutic impact on cells, as cells might be fully damaged, i.e., all DNA binding sites are occupied by the drug.

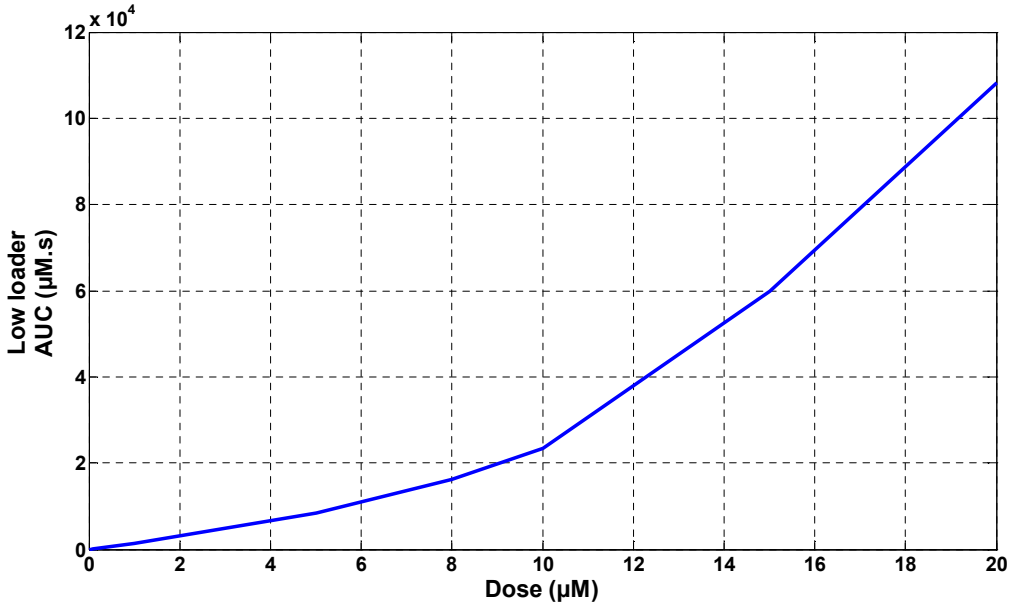


Figure 4.27: AUC for L_n over 3600 seconds (1 hour) against dose D using the parameter values for the low loader.

4.7 Summary

In summary, this chapter presents a state-space model for the *in vitro* uptake kinetics of the anti-cancer agent TPT has been extended from a previously published model by Evans *et al.* [13]. The extended model describes drug activity and delivery of the pharmacologically active form to the DNA target as well as the catalysis of the ALDH enzyme and the elimination of drug from the cytoplasm via the efflux pump. Structural identifiability analysis of the cell-based model was performed using the software tool MATHEMATICA which demonstrates that all of the unknown model parameters are uniquely determined by the output structure corresponding to the real experiment. Parameter estimation was performed using the software tool FACSIMILE. Model simulations have been compared with MCF-7 cells and found to give good qualitative agreement. This model will be used in the drug kinetics/dynamics model in Chapter 6.

5 The cell cycle model

5.1 Introduction

In recent years, molecular biologists have revealed a wealth of information about the proteins regulating and controlling cell growth and division (i.e., the cell cycle and mitosis) in eukaryotic cells [157]. Moreover, the dynamic properties of this complex regulatory system defy intuitive reasoning alone. Quantitative tools using software packages are necessary to probe, understand and reliably predict the behaviour of such networks (i.e. cell cycle control). In this thesis, the hypothetical molecular mechanisms (molecular wiring diagrams) of the cell cycle are converted into sets of non-linear ODEs in computer-readable form [207]. Numerical solution of the ODEs can then be compared with experimental results (observations) to determine unknown model parameters (including kinetic rate constants) and to confirm the adequacy of the mathematical model. The mathematical model describing the cell cycle events, with mechanistically meaningful equations and rate constants, is derived with the intention of giving precise simulations of known observations (experimental results). Such models once validated may also allow predictions about the behaviour of the system that can be verified experimentally.

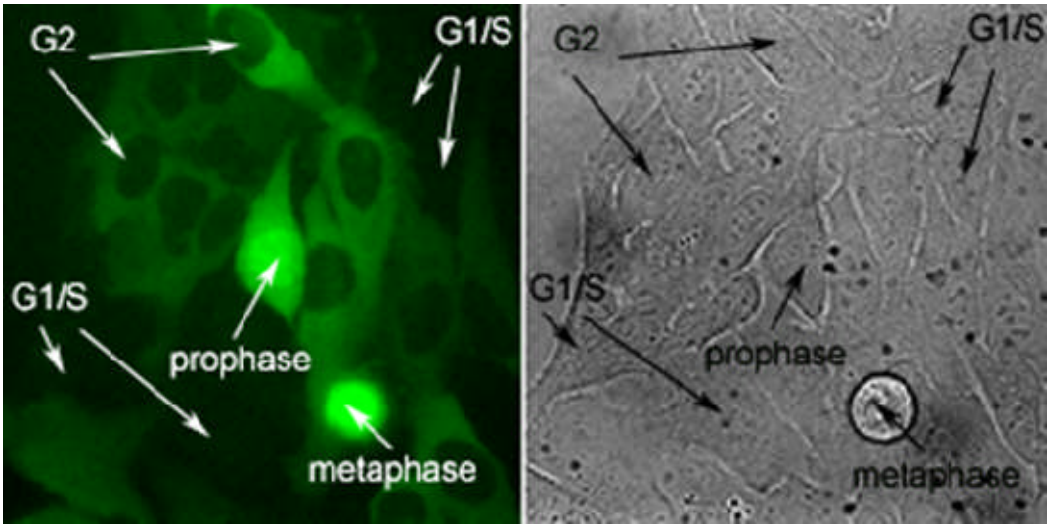
This chapter presents a mathematical model describing the growth of single human cells in the absence of the anti-cancer agent TPT. The cell cycle model described in this thesis is a novel extension of a previously published model by Pomerening *et al.* [14] (for *Xenopus laevis* embryos) to account for the inhibitory action of the protein $p21^{CIP1/WAF1}$ (under the transcriptional control of p53, see Section 3.3) on Cyclin B1. The extended model is based on underlying biological assumptions of the cell cycle process, and the unknown model parameters of the proposed model are estimated from high-resolution fluorescence of Cyclin B1 data (see Section 5.2) from live human osteosarcoma cells (U-2 OS cell line) provided by the project collaborators at Cardiff University. The analytical and numerical methods described in Chapter 2 are used to study the cell cycle process presented in this work.

5.2 Cyclin B1-GFP experiments

Critical to modelling the sequence of events of the cell cycle process is the ability to undertake high spatio-temporal resolution monitoring of the cell cycle progression using green fluorescent protein (GFP) cell cycle phase markers (CCPMs). This monitoring technique allows the continuous tracking of single cell checkpoint transitions within a heterogeneous population in a non-destructive (non-invasive) manner. The GFP-based probe has expression, localisation and destruction characteristics that shadow the dynamic behaviour of the Cyclin B1 regulator (during the cell cycle) in living cells [159]. That is, the status of the cell cycle is monitored and reported by dynamic stealth sensors based on the expression, localisation and degradation of GFP under the control of one of the key components of the cell cycle regulators, the Cyclin B1 protein. These sensors can precisely and non-invasively determine the status of the cell cycle in living cells by fluorescence imaging, i.e., without perturbing the progression of the cell cycle. The performance of the non-invasive stealth reporter has been validated on high content to high throughput detection platforms including multi-well high-throughput screen (HTS) imaging, single cell kinetic-tracking and multi-parameter flow cytometry [107, 159, 208]. HTS is a method of scientific experimentation used particularly in the field of drug discovery and the related fields of biology and chemistry. This method allows researchers to rapidly conduct millions of biochemical and pharmacological tests.

Cyclin B1-GFP tracking provides important sub-phase information about the progression of the cell cycle, the dynamics of the cell-cycle regulator (Cyclin B1) in parallel with morphological landmarks and DNA content analysis (Figure 5.1). Moreover, this time-lapse microscopy provides a means for understanding the subtleties in cell cycle event patterns such as delayed versus arrested patterns [209, 210]. The continuous progression of the cell cycle traverse and encoded molecular readouts in bifurcation cell lineages have been tracked to enable data extraction with subsequent linking to phenotypic cellular behaviour in both normal conditions (i.e., in the absence of TPT), and in response to treatment with the anti-cancer drug TPT (Chapter 6). In experimental terms, a cell lineage is defined as a descent in a line from a mutual progenitor (origin or starting cell) that was exposed to a given influence for a discrete time period [211]. The behaviour of both the

progenitor and the descending line of offspring (descendants) demonstrates the time-integrated cell cycle response (in the absence of the anti-cancer drug TPT) and dynamic response (in the presence of TPT) [210, 212]. Examples of such responses include changes in the inter-mitotic time (IMT), i.e., the temporal distance between two consecutive cell divisions, as well as cell death [211]. The experimental data provide the essential data upon which the cell cycle model is both developed and validated to a certain degree.



(a) **(b)**

Figure 5.1: Human osteosarcoma cell (U-2 OS cells): (a) Expressing a Cyclin B1-GFP stealth reporter. (b) A corresponding transmission image to identify all the cells in the field of view (taken from Reference [107]).

The parental cell line used in this thesis was a human osteosarcoma cell line (U-2 OS) derived from a 15 year old Caucasian female. U-2 OS cells were transfected (by the project collaborators at Cardiff University) with G_2M CCPM (GE Healthcare, UK) using the transfection reagent Fugene (Roche). Tracking cell cycle progression using the high-resolution fluorescence microscopy was performed with cells seeded into a multi-well coverslip-bottomed plate. Each well represents a different treatment regimen (a dose range for the anti-cancer agent TPT of 1-10 μM, see Section 6.2) in addition to a control well (untreated cells). Cyclin B1-GFP data used for estimating the unknown parameters of the extended cell cycle model (see Sections 5.3, 5.4 and 5.5) are for untreated single human osteosarcoma cells. The cultured dishes were placed onto a time-lapse instrument designed to capture bright-field phase images and Cyclin B1-GFP fluorescence images. Sequences of images were captured every 20 minutes for 48 hours (h), in triplicate, i.e. three regions of interest (ROI's), in the control experiment.

For the cell of interest, in each image frame, three ROIs are used to extract parameters from the raw image sequence (using FLUROTRAK, a bespoke software tool by the Cardiff collaborators) viewed in the image analysis software METAMORPH. The first ROI is positioned on the nucleus and the other two ROIs are positioned on the cytoplasmic regions (usually, on the opposite sides of the nucleus of the cell of interest). Once the ROIs are placed, the cell of interest is tracked frame by frame for 48 h (resulting in a total of 144 frames) starting at frame one. A normal cell division results in two daughter cells, however in some cases, a progenitor cell can abnormally generate three or four individual daughter cells, i.e., cells can produce re-fused and/or polyploidy (when cells have more than two pairs of chromosomes) outcomes. The extended cell cycle model in this thesis looks into cells that experience normal cytokinesis (see Section 3.3) and such outcomes (i.e. re-fused or polyploidy) are not in the scope of this work.

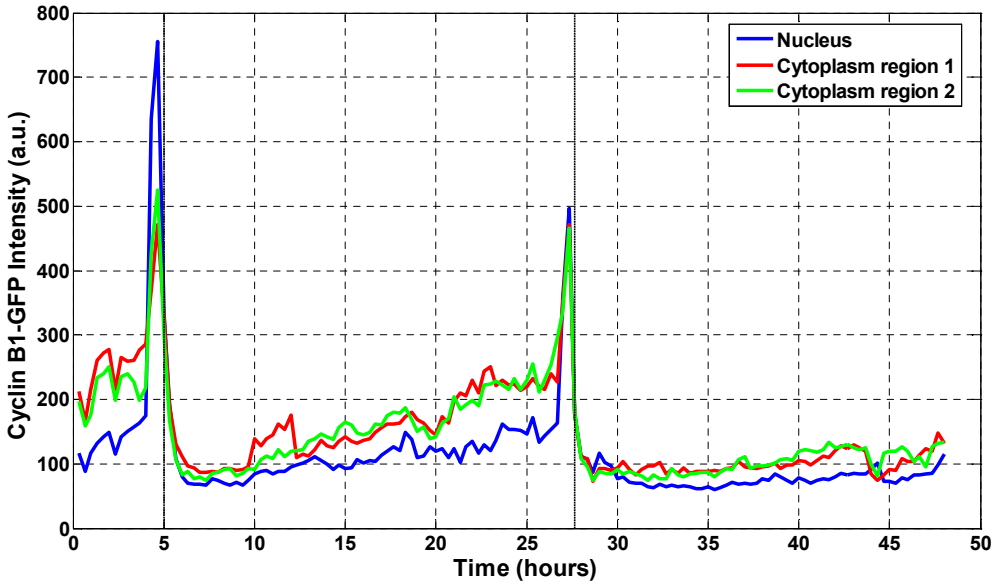


Figure 5.2: Continuous Cyclin B1-GFP intensity profile track extracted from an untreated G_2 cell, three compartments were tracked, the nucleus (blue) and two regions of interest in the cytoplasm (red and green) respectively.

Continuous cell cycle tracking at the single cell level (i.e., a typical time-lapse microscopy image sequence) shows cells traversing the cell cycle and fluorescence changes (see Figure 5.2) as the cell progresses to M phase from G_1. Individual cells rapidly increase Cyclin B1-GFP expression (i.e. become brighter) and a translocation event from the cytoplasm to the nucleus compartment occurs (prior to mitosis) via the cytoplasmic retention signal [213]. That is, the fluorescent G_2M CCPM reporter system relies on the control of location and

levels of expression of GFP (driven by the promoter region and removal via the destruction box) as cells progress to the later phases of the cell cycle and negotiate entry and exit to the M phase. This is attained by utilising the functional components from the Cyclin B1 protein to confer switch-like properties to the stealth reporter. Expression levels of Cyclin B1 are tightly regulated and act as a control switch that is appropriate for transition from S phase through G_2 into mitosis (and cytokinesis). Due to the absence of the cyclin box from the reporter, the stealth sensor does not interfere or perturb the progression of the cell cycle.

A typical cell lineage over 48 h illustrates cellular information at two levels: (a) phenotypic behaviour (cell division represented by M) and (b) Cyclin B1-GFP reporter readout, i.e., fluorescence intensities in arbitrary units (a.u.) and hence, position in the cell cycle process at the single cell level as shown in Figure 5.2. The continuous progression of the cell cycle traverse between the two landmarks, represented by the mitotic events (M_1 and M_2) at both ends was demonstrated for a typical cell in the G_2 phase of the cell cycle.

5.3 Mathematical model of the cell cycle

Most molecular elements are affected by the cell cycle, however, this process is controlled by a very limited set of biochemical interactions containing a few tens of components [214]. Some of these interactions are essential, while preventing other biochemical interactions does not prohibit the cell cycle and division process [215]. The cell cycle process happens as a result of external signalling, however, in the case of tumours, this tight regulation is absent (both intra- and extracellular) [215], and an internal set of protein-protein interactions and genetic regulation can result in unconstrained cell cycle division [216]. Accordingly, it can be assumed that the oscillations in the cell cycle process can be ascribed to a simple oscillatory core which has also been named an *oscillophore* by Goldstein *et al.* [217]. Such unlimited cell cycles that are not controlled by external stimulations and/or feedback mechanisms from the size of the cell are frequent [163, 218-220].

5.3.1 Cell cycle models

One of the major challenges for theoretical molecular biologists is to describe the physiology of cell proliferation in a wide assortment of unicellular and multi-cellular eukaryotes in terms of their underlying molecular control systems. Therefore, such molecular connections are necessary which will be made by developing mathematical (computational) models that reflect the complexity of real cell cycle controls [221, 222]. In order to design such models and understand how they function, the basic control mechanisms should be well understood.

Within the field of cancer research, significant efforts have been devoted to modelling the cell cycle process [14, 157, 221, 223-242]. The large number of such mathematical models, often based on different components and different feedback loops, indicates that the same biological response can be described using different mathematical techniques [243]. Accordingly, a criterion for selecting an appropriate mathematical description is required. Most existing ODE-based cell cycle models are based on a hysteresis mechanism, such as the mathematical models of the cell cycle process developed by Tyson and Novak [157] (for yeast cells) and Pomerening *et al.* [14, 244], or on an oscillophore which was proposed by Goldbeter [224], and later adopted by several research groups including Hatzimanikatis *et al.* [223] as well as Tyson and Novak [157], each with different components, but having the same structure [243]. In the first group of cell cycle models, i.e., the models based on a hysteresis mechanism, the internal parameters are of no vital importance, since the external parameters determine both the presence and the length of the cell cycle. However, oscillophores can be sensitive to parameter changes. The paper by Clyde *et al.* [245] is an excellent review on the development of cell cycle models.

5.3.2 The extended cell cycle model

The cell cycle model described in this thesis is an extended version of the model developed by Pomerening *et al.* [14] to account for the inhibition dynamics of $p21^{CIP1/WAF1}$ protein on Cyclin B1 for mammalian cells. In this model, biochemical mechanisms are translated into

systems of non-linear ODEs by the law of mass action. That is the mathematical modelling of the cell cycle process involves the construction of ODEs to describe protein interaction networks that regulate the cell cycle process. Moreover, dynamical systems theory is used to analyse the qualitative behaviour of these equations and to bridge the gap between mechanisms and theoretical molecular biology. The original cell cycle model [14] relies on a hysteresis mechanism, and therefore, the choice was made on this set of models since the cell cycle may vary in length as well as in the concentration of its different components, and the presence of an oscillophore appears not to be sensitive to the precise properties of the elements used to steer the cell cycle [243].

Nurse [246] proposed that the DNA replication-division cycle in all eukaryotic cells is regulated by a common set of protein interaction networks (i.e. the key elements and biochemical interactions are similar in eukaryotes), which has been confirmed by numerous experimental studies [214, 247]. However, the mix of these proteins and interactions varies among organisms and therefore, each organism generates its own idiosyncrasies of cell growth and division. The generic features of the cell cycle control involve these common genes and regulatory proteins in addition to the general dynamic principles by which they coordinate the replication and division of the genome from the mother cell to offspring (two daughter cells). The peculiarities of the cell cycle process concern exactly which parts of the common mechanisms are functioning in any given cell type, given the genetic background and developmental stage of an organism [247]. Therefore, given the data available (i.e., Cyclin B1-GFP data, see Figure 5.2), the cell cycle model in this thesis is adopted from the mathematical model in Pomerening *et al.* [14] which was originally derived for the cell cycle of *Xenopus laevis* (African clawed frog) embryos to model the cell cycle control systems in mammalian cells.

The original model of Pomerening *et al.* [14] is similar in many ways to that of Tyson and Novak [157] (for yeast cells): both rely on a hysteresis mechanism and the governing equations account for the regulation of CDK1/Cyclin B1. When comparing the simulation results of the governing equations of Pomerening *et al.* [14] in Figure 5.3 and Tyson and Novak [157] in Figure 5.4, the original cell cycle model by Pomerening *et al.* [14] reproduces closer qualitative agreement to experimentally observed results in Figure 5.5 (i.e., Cyclin B1-GFP experiments). Accordingly, the choice has been made to extend the

previously published cell cycle model by Pomerening *et al.* [14] to mathematically describe the cell cycle process to account for the effects of CKI $p21^{CIP1/WAF1}$, using Cyclin B1-GFP data from high-resolution fluorescence microscopy experiments for live human osteosarcoma cells.

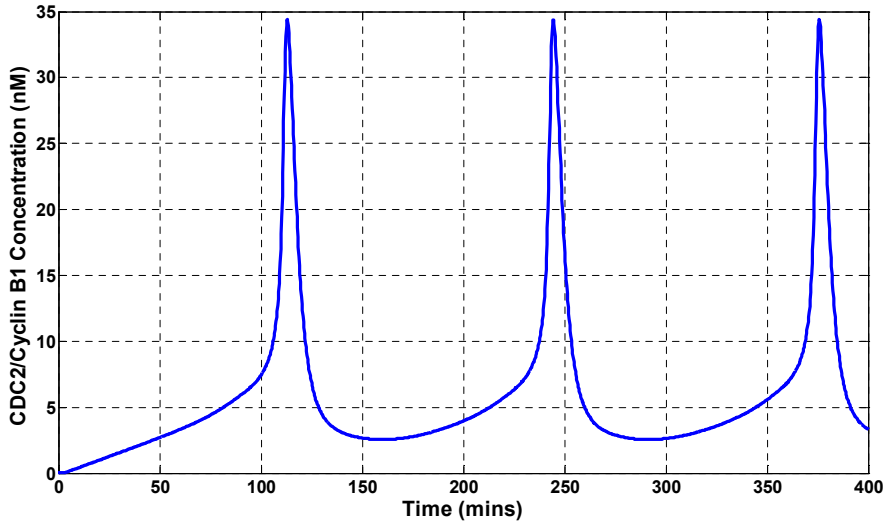


Figure 5.3: Simulation of CDC2/Cyclin B1 activity using the cell cycle model of Pomerening *et al.* [14] for *Xenopus laevis* embryos.

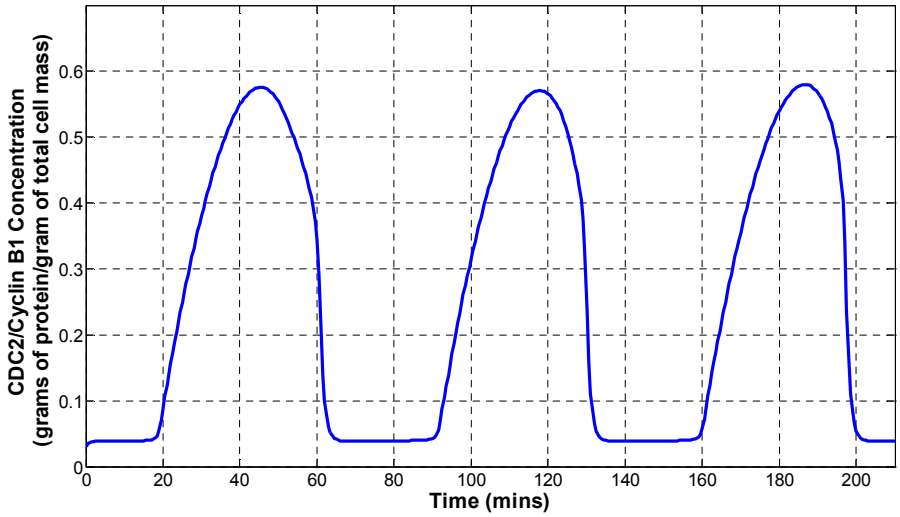


Figure 5.4: Simulation of CDK1/Cyclin B1 activity using the cell cycle model of Tyson and Novak [157] for yeast cells.

The cell cycle process of the *Xenopus laevis* embryo is a prominent example of a robust biological oscillator. In addition, the cycles of CDC2 (or CDK1) activation and inactivation are quasi-synchronous and rapid with a limited variability among embryos

[244]. The cycles of the *Xenopus laevis* embryo continue with relatively normal timing and amplitude even under conditions when the embryos are enucleated [248] or when cell-cycle endpoints (such as DNA replication and spindle assembly) are met. According to the observations, the embryonic cell cycle is driven by an autonomous, modular, clock-like regulatory circuit centred on CDC2 (or CDK1) and the APC [249, 250].

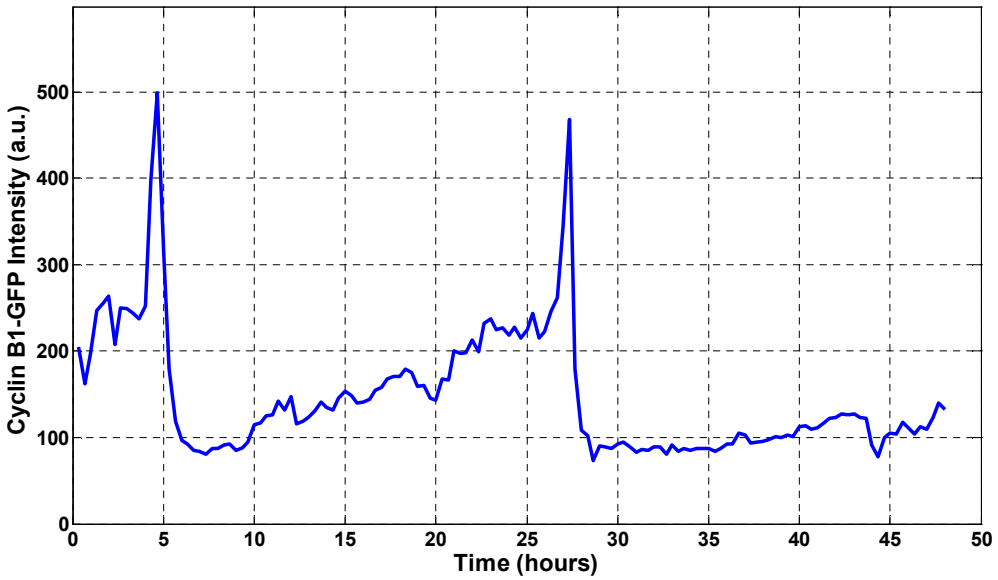


Figure 5.5: Average cytoplasmic Cyclin B1-GFP intensity profile track of the untreated G_2 cell in Figure 5.2.

The original cell cycle model by Pomerening *et al.* [14] is based on a simplified model for the CDC2 (or CDK1)/APC system (see Section 3.3), as shown in Figure 5.6. Synthesis and gradual accumulation of the major mitotic cyclin (Cyclin B1 protein) drives the cell into mitosis. This accumulation results in the binding of Cyclin B1 to CDK1 (CDC2), and, under appropriate conditions, the CDK1/Cyclin B1 dimer becomes active and phosphorylates mitotic substrates, hence driving the G_1/S transition (*Start*), see Section 3.3. The transition from M phase back to G_1 (*Finish*) is driven by a negative-feedback loop; the active CDK1/Cyclin B1 dimer activates the APC, resulting in polyubiquitination[7] [251] and proteolysis (degradation of proteins) of Cyclin B1 [252-254]. The presence of a negative-feedback loop does not guarantee persistent oscillations of the CDC2/APC network. The importance of this negative-feedback loop for mitotic oscillations was demonstrated experimentally when APC-resistant cyclin mutants were added to *Xenopus*

[7] Polyubiquitination: modification of a protein by the addition of one or more ubiquitin monomers. Ubiquitins are polypeptides involved in modifying and degrading proteins.

egg extracts in the G_1 phase (interphase), the extracts approached a steady state rather than oscillating [255].

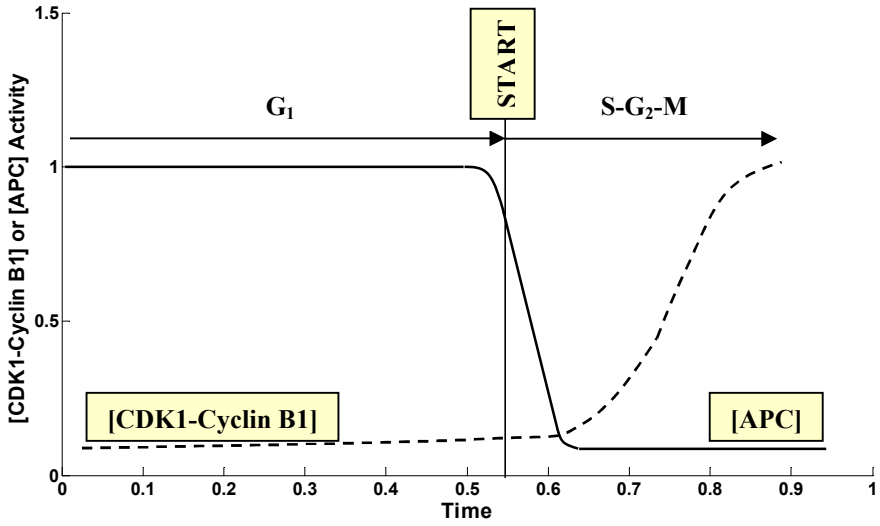


Figure 5.6: At G_1, APC negates the activity of CDC2 (or CDK1) by degrading the major mitotic cyclin (Cyclin B1), while the activity of APC is inhibited by the CDK1/Cyclin B1 dimer (redrawn from Reference [107]).

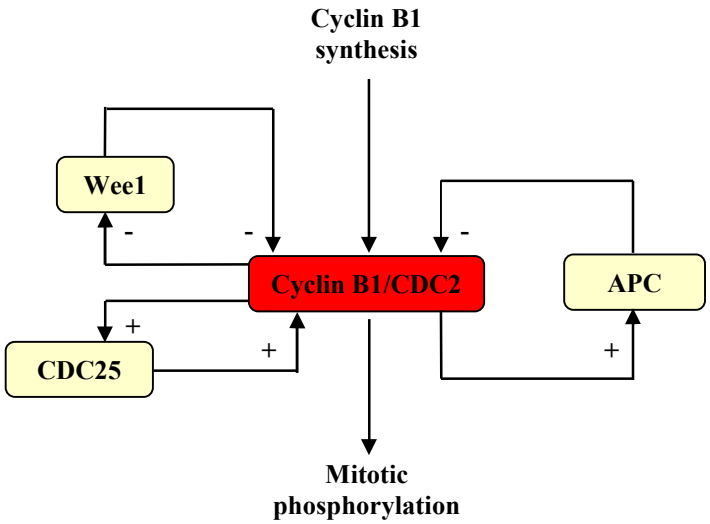


Figure 5.7: Schematic depiction of the CDC2/APC network by Pomerening *et al.* [14].

A long negative-feedback loop (in some cases) by itself is sufficient to produce oscillations. However, the CDC2/APC network in Pomerening *et al.* [14] contains two additional elements (i.e., in addition to the long negative-feedback loop) in the circuit presented in Figure 5.7: a double-negative-feedback loop and a positive-feedback loop centred on Cyclin B1/CDC2 [14, 244], as shown in Figure 5.7. That is, active Cyclin B1/CDC2 phosphorylates and activates its activator (represented by the positive-feedback

loop), the phosphatase CDC25 (Figure 3.10, Section 3.3). Moreover, active Cyclin B1/CDC2 phosphorylates and inactivates its inhibitors (represented by double-negative-feedback loop), Wee1 and Myt1 protein kinases (Figure 3.10, Section 3.3).

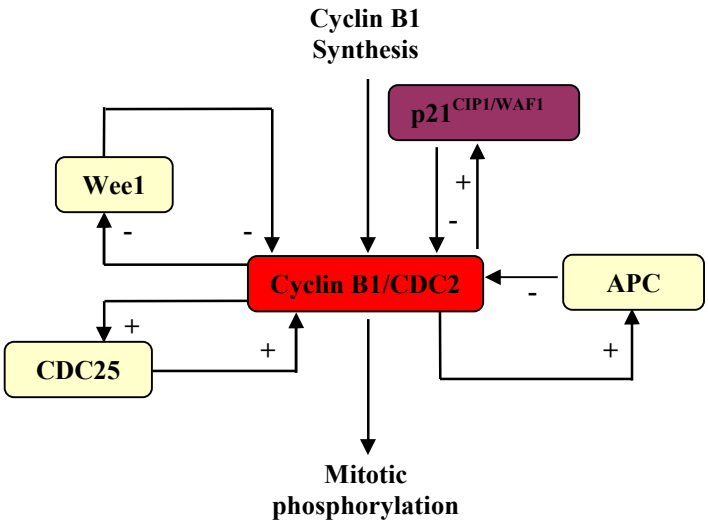


Figure 5.8: Schematic depiction of the CDC2/APC network in the extended model to contain an additional circuit element, a negative-feedback loop to account for the inhibitory action of $p21^{CIP1/WAF1}$.

In this thesis, the original cell cycle model by Pomerening *et al.* [14] has been extended to contain an additional circuit element, a negative-feedback loop, to account for the inhibitory action of $p21^{CIP1/WAF1}$ as shown in Figure 5.8. The CKI $p21^{CIP1/WAF1}$ and Cyclin B1/CDC2 are antagonist proteins, in fact, tumour cells induce the accumulation of this CKI under the transcriptional control of p53. The accumulating $p21^{CIP1/WAF1}$ binds to the Cyclin B1/CDC2 dimer and therefore inhibiting its activity for initiating DNA synthesis (see Section 3.3). This results in stabilising the G_1 state (G_1 arrest in some cases) of the cell cycle process. These loops could function as a bistable system (switch) [14, 221, 236, 244] with a hysteretic steady-state stimulus/response relationship [256], toggling between two discrete steady states, namely, the G_1 state and the S-G_2-M state. A system is bistable if it can switch between two discrete alternative states without being able to rest in an intermediate state (Figure 5.9). In the first case (i.e., G_1 phase), CDC2 and CDC25 are inactive while Wee1, Myt1, APC and $p21^{CIP1/WAF1}$ are active and in the second stable state (i.e., S-G_2-M state), CDC2 and CDC25 are active while Wee1, Myt1, APC and $p21^{CIP1/WAF1}$ are inactive.

A bistable trigger in the extended cell cycle model and the original model (Figures 5.7 and 5.8, respectively) is critical for the function of a mitotic oscillator by ensuring that a cell settles in a discrete, mutually exclusive G_1 phase and S-G_2-M state and not in a continuum of intermediate states [221, 257]. The system will approach a high activity (S-G_2-M state) of CDC2 above a critical concentration of non-degradable Cyclin B1 and below some other lower concentration of Cyclin B1, the system will approach a low CDC2 activity (G_1 state). At an intermediate concentration of Cyclin B1, the system will approach either the G_1 state or S-G_2-M state depending upon the initial conditions, but the system does not settle in a state with a partially activated pool of CDC2. This property (i.e., bistability of CDC2 activation) minimises chattering (slipping rapidly back and forth between cell cycle phases) during the transitions into and out of S-G_2-M. The advantage of the bistability property is to ensure that the mitotic oscillator will never approach a stable steady-state. Adding a bistable trigger to the negative-feedback loop (i.e., Cyclin B1 destruction and CDC2 inactivation by APC) as suggested by Pomerening *et al.* [14] in Figure 5.7, in addition to the second negative-feedback loop (the inhibition effect by the $p21^{CIP1/WAF1}$ protein) in the extended model (Figure 5.8) also produces sustained oscillations with explosive spikes of CDC2 activity, resulting in a relaxation oscillator. This category of oscillators is suitable as a biological timer [14] and has advantages over simple negative-feedback oscillators in terms of noise rejection, reliability, spatial propagation and self-synchronisation [258-260].

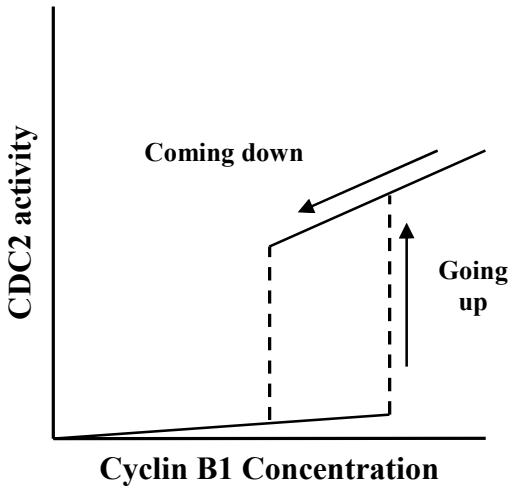


Figure 5.9: The bistable response could arise from a combination of ultra-sensitivity and positive feedback (redrawn from Reference [14]).

Pomerening *et al.* [14] have experimentally determined that the CDC2 system is bistable. That is, the response of CDC2 to active Cyclin B1 protein is temporally abrupt and switch-like. Moreover, it has been shown in the study by Pomerening *et al.* [14], that the activation of CDC2 exhibits hysteresis. The Cyclin B1-GFP data shown in Figure 5.10 (see Section 5.2) reflect CDC2 activity (Cyclin B1-GFP fluorescence intensities) of two different daughter cells (labelled Track 1 and Track 3) from the same progenitor cell. The time course of the activation of CDC2 in response to Cyclin B1 exhibits a temporal lag that precedes abrupt transition (G_2/M) between low and high Cyclin B1-GFP intensity (Figure 5.10), i.e., CDC2 activation. This means that the rate of CDC2 activation increases with CDC2 activity. Therefore, the activation of CDC2 is bistable, this observed temporal abruptness (highly switch-like with explosive spikes) is consistent with the predicted behaviour of bistable systems. Another feature of bistable systems is the discontinuity in the steady-state level of CDC2 activity as a function of Cyclin B1 concentration. These observations in Figure 5.10 are consistent with a bistable response, that is, small changes in Cyclin B1 result in large changes in CDC2 activity.

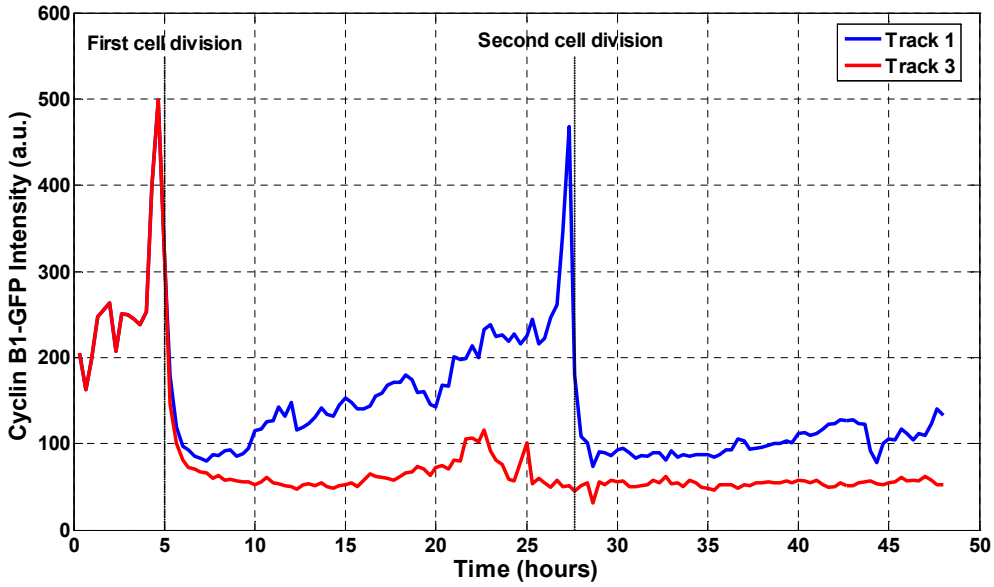


Figure 5.10: Average cytoplasmic Cyclin B1-GFP intensity profile along two different tracks (Track 1 and Track 3) resulting from the same progenitor (untreated) cell of the exemplar lineage (see Figure 5.2).

Temporal abruptness (between low and high CDC2 activity) and discontinuity in the steady-state level of CDC2 activity (as function of Cyclin B1 concentration) can also be detected in a highly switch-like, ultrasensitive, monostable system. A more definitive

approach to distinguish between monostable and bistable responses is to examine hysteresis in the response of CDC2 to Cyclin B1. A monostable system will approach the steady-state level of CDC2 activity for a given fixed Cyclin B1 concentration, in spite of the history of the system. However, a bistable system will approach either of the two stable steady states, namely, G_1 state and S-G_2-M state, depending on the history of the system. Accordingly, as shown in Figure 5.9, the stimulus-response curve for a bistable system should exhibit hysteresis, with a relatively low threshold for switching to the G_1 state and with a relatively higher threshold for switching to the S-G_2-M state. The CDC2 activity is bistable, the concentration of Cyclin B1 in Track 3 following the first division (Figure 5.10) is not sufficient to drive the cell from the G_1 phase to the S-G_2-M state as a result of the inhibition of the CKI $p21^{CIP1/WAF1}$ (G_1 arrest). However, the concentration of Cyclin B1 in Track 1 following the first division is sufficient to maintain an S-G_2-M state for long periods of time resulting in another cytokinesis (second cell division) before switching again to the G_1 state. Therefore, the Cyclin B1-GFP data confirm that the CDC2 activation exhibits hysteresis, which is consistent with the predicted behaviour of the cell cycle process, i.e., the mitotic trigger is bistable.

Figure 5.11 provides a schematic diagram of the structure of the extended cell cycle model of the CDC2/APC system including the effects of the CKI $p21^{CIP1/WAF1}$. The biochemical reactions (processes) of the postulated mathematical model are described by the following system of non-linear ODEs:

$$\begin{aligned}\frac{\mathrm{d[CycB]}}{\mathrm{dt}} &= C_1 k_{synth} - C_1 k_{dest}([\mathrm{APC_{active}}]+[\mathrm{p21_{active}}])[\mathrm{CycB}] \\ &\quad - C_1 k_a[\mathrm{CDC2}][\mathrm{CycB}] + C_1 k_d[\mathrm{CDC2CycB}]\end{aligned} \tag{5.1}$$

$$\begin{aligned}\frac{\mathrm{d[CDC2CycB]}}{\mathrm{dt}} &= C_1 k_a[\mathrm{CDC2}][\mathrm{CycB}] - C_1 k_d[\mathrm{CDC2CycB}] \\ &\quad - C_1 k_{dest}([\mathrm{APC_{active}}]+[\mathrm{p21_{active}}])[\mathrm{CDC2CycB}] \\ &\quad - C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB}] \\ &\quad - C_1 C_2 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{total}}][\mathrm{CDC2CycB}] \\ &\quad + C_1 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB}] \\ &\quad + C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YP}}] \\ &\quad + C_1 C_2 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{total}}][\mathrm{CDC2CycB_{YP}}] \\ &\quad - C_1 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YP}}]\end{aligned} \tag{5.2}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2CycB_{YP}}]}{\mathrm{dt}} &= C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB}] \\ &- C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YP}}] \\ &- C_1 k_{\mathrm{CAK}}[\mathrm{CDC2CycB_{YP}}] + C_1 k_{\mathrm{PP2C}}[\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{dest}([\mathrm{APC_{active}}] + [\mathrm{p21_{active}}])[\mathrm{CDC2CycB_{YP}}] \\ &+ C_1 C_2 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{total}}][\mathrm{CDC2CycB}] \\ &- C_1 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB}] \\ &- C_1 C_2 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{total}}][\mathrm{CDC2CycB_{YP}}] \\ &+ C_1 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YP}}]\end{aligned} \tag{5.3}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2CycB_{YPTP}}]}{\mathrm{dt}} &= C_1 k_{\mathrm{CAK}}[\mathrm{CDC2CycB_{YP}}] - C_1 k_{\mathrm{PP2C}}[\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}] \\ &+ C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}] \\ &- C_1 k_{dest}([\mathrm{APC_{active}}] + [\mathrm{p21_{active}}])[\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 C_2 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{total}}][\mathrm{CDC2CycB_{YPTP}}] \\ &+ C_1 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}] \\ &+ C_1 C_2 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{total}}][\mathrm{CDC2CycB_{TP}}] \\ &- C_1 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}]\end{aligned} \tag{5.4}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2CycB_{TP}}]}{\mathrm{dt}} &= C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{dest}([\mathrm{APC_{active}}] + [\mathrm{p21_{active}}])[\mathrm{CDC2CycB_{YPTP}}] \\ &+ C_1 C_2 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{total}}][\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{\mathrm{CDC25basal}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}] \\ &- C_1 C_2 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{total}}][\mathrm{CDC2CycB_{TP}}] \\ &+ C_1 k_{\mathrm{Wee1basal}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}]\end{aligned} \tag{5.5}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2}]}{\mathrm{dt}} &= C_1 k_{dest}([\mathrm{APC_{active}}] + [\mathrm{p21_{active}}])([\mathrm{CDC2CycB}] \\ &+ [\mathrm{CDC2CycB_{YP}}] + [\mathrm{CDC2CycB_{YPTP}}] \\ &+ [\mathrm{CDC2CycB_{TP}}]) \\ &- C_1 k_a[\mathrm{CDC2}][\mathrm{CycB}] + C_1 k_d[\mathrm{CDC2CycB}]\end{aligned} \tag{5.6}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{APC_{active}}]}{\mathrm{dt}} &= \frac{C_1 C_2 k_{\mathrm{APCon}}[\mathrm{APC_{total}}][\mathrm{CDC2CycB_{TP}}]^{n_{\mathrm{APC}}}}{(C_2[\mathrm{EC50_{APC}}])^{n_{\mathrm{APC}}} + [\mathrm{CDC2CycB_{TP}}]^{n_{\mathrm{APC}}}} \\ &\frac{-C_1 k_{\mathrm{APCon}}[\mathrm{APC_{active}}][\mathrm{CDC2CycB_{TP}}]^{n_{\mathrm{APC}}}}{(C_2[\mathrm{EC50_{APC}}])^{n_{\mathrm{APC}}} + [\mathrm{CDC2CycB_{TP}}]^{n_{\mathrm{APC}}}} \\ &- C_1 k_{\mathrm{APCoff}}[\mathrm{APC_{active}}]\end{aligned} \tag{5.7}$$

$$\frac{\mathrm{d}[\mathrm{Wee1}_{\mathrm{active}}]}{\mathrm{dt}} = \frac{-C_1 k_{\mathrm{Wee1off}}[\mathrm{Wee1}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{Wee1}}}}{(C_2[\mathrm{EC50}_{\mathrm{Wee1}}])^{n_{\mathrm{Wee1}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{Wee1}}}} + C_1 k_{\mathrm{Wee1on}}(C_2[\mathrm{Wee1}_{\mathrm{total}}] - [\mathrm{Wee1}_{\mathrm{active}}]) \tag{5.8}$$

$$\frac{\mathrm{d}[\mathrm{CDC25}_{\mathrm{active}}]}{\mathrm{dt}} = \frac{C_1 C_2 k_{\mathrm{CDC25on}}[\mathrm{CDC25}_{\mathrm{total}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}}{(C_2[\mathrm{EC50}_{\mathrm{CDC25}}])^{n_{\mathrm{CDC25}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}} - \frac{C_1 k_{\mathrm{CDC25on}}[\mathrm{CDC25}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}}{(C_2[\mathrm{EC50}_{\mathrm{CDC25}}])^{n_{\mathrm{CDC25}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}} - C_1 k_{\mathrm{CDC25off}}[\mathrm{CDC25}_{\mathrm{active}}] \tag{5.9}$$

$$\frac{\mathrm{d}[\mathrm{p21}_{\mathrm{active}}]}{\mathrm{dt}} = \frac{C_1 C_2 k_{\mathrm{p21on}}[\mathrm{p21}_{\mathrm{total}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}}{(C_2[\mathrm{EC50}_{\mathrm{p21}}])^{n_{\mathrm{p21}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}} - \frac{C_1 k_{\mathrm{p21on}}[\mathrm{p21}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}}{(C_2[\mathrm{EC50}_{\mathrm{p21}}])^{n_{\mathrm{p21}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}} - C_1 k_{\mathrm{p21off}}[\mathrm{p21}_{\mathrm{active}}] \tag{5.10}$$

where the components in the square brackets refer to concentrations. In the extended model of the CDC/APC system described by Equations (5.1)–(5.10), there are four different forms of the Cyclin B1/CDC2 dimer: [CDC2CycB], [CDC2CycB$_{\mathrm{YP}}$], [CDC2CycB$_{\mathrm{YPTP}}$] and [CDC2CycB$_{\mathrm{TP}}$] denote the concentrations (in nM) of the Cyclin B1/CDC2 dimer forms, namely, non-phosphorylated complex (inactive), Y15-phosphorylated complex (inactive), Y15- and T161-phosphorylated complex (inactive) and T161-phosphorylated (active), respectively, at time *t* of the Cyclin B1-GFP tracking. The time-lapse microscopy experiments provide intensity profiles of the active form of Cyclin B1/CDC2 complex (T161-phophorylated complex) in the nucleus and cytoplasm compartments of human osteosarcoma cells (U-2 OS cell line).

In the extended model, [CycB] and [CDC] are the concentrations (in nM) of free (unbound) monomers of Cyclin B1 and CDC2 (CDK1), respectively. The APC, the protein kinase Wee1, the phosphatase CDC25, and the CKI p21$^{\mathrm{CIP1/WAF1}}$ all exist in, active and inactive forms (Figure 5.11). In the mathematical model described by Equations (5.1)-(5.10), [APC$_{\mathrm{active}}$], [Wee1$_{\mathrm{active}}$], [CDC25$_{\mathrm{active}}$] and [p21$_{\mathrm{active}}$] denote the concentrations (in nM) of the active forms of APC, Wee1 protein kinase, CDC25 phosphatase and p21$^{\mathrm{CIP1/WAF1}}$, respectively. In addition, their total concentrations (in nM) are denoted by [APC$_{\mathrm{total}}$], [Wee1$_{\mathrm{total}}$], [CDC25$_{\mathrm{total}}$] and [p21$_{\mathrm{total}}$], respectively.

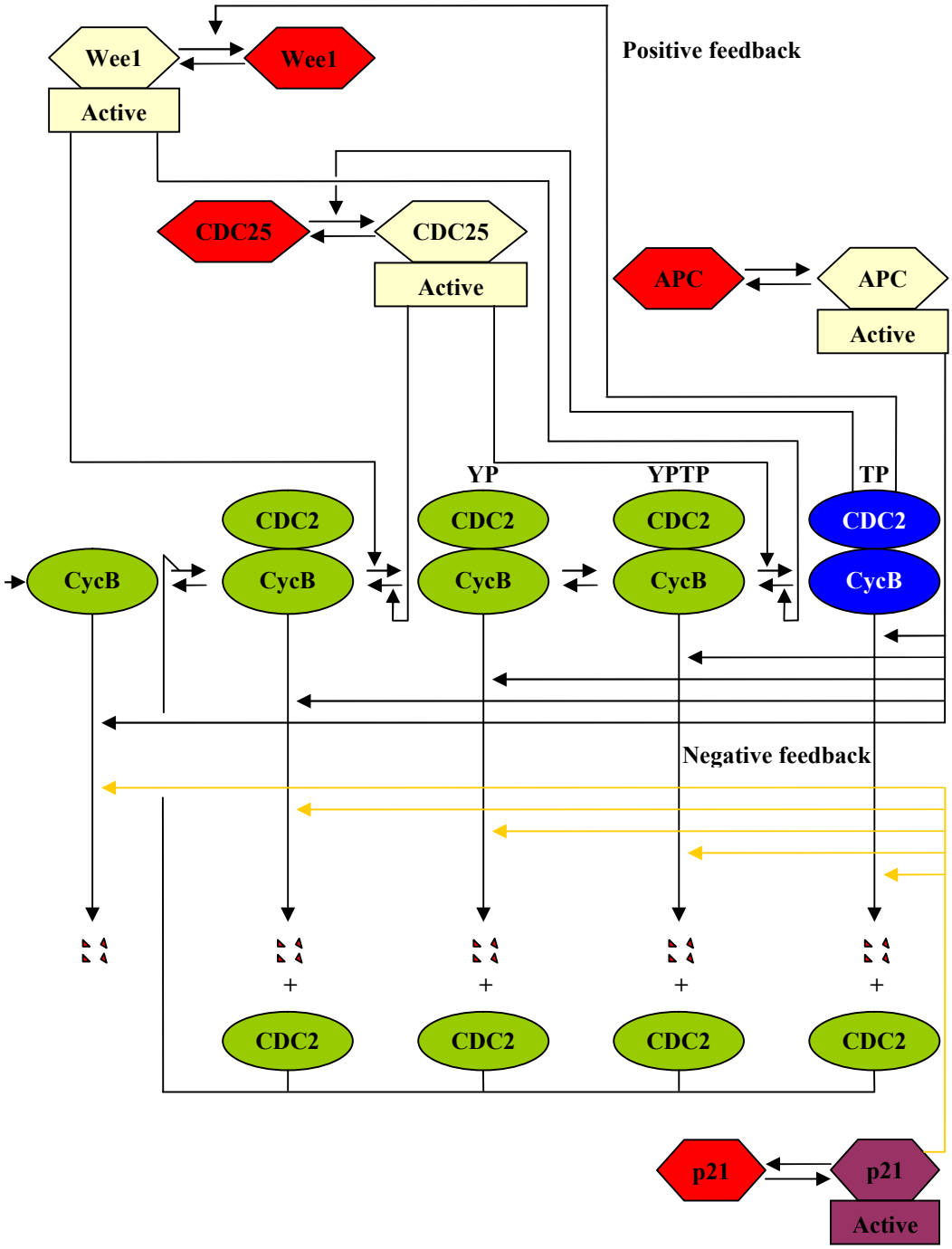


Figure 5.11: Schematic diagram of the extended CDC2/APC model to account for the effects of the CKI p21$^{CIP1/WAF1}$.

In the extended model, the rate at which Cyclin B1 monomer ([CycB]) is synthesised is modelled via zero-order chemical reaction with constant k_{synth}. Additionally, the rate at which Cyclin B1 forms (monomer and dimers) are degraded is assumed to be proportional (with the rate constant k_{dest}) to the product of the sum of concentrations of the active forms of APC and p21$^{CIP1/WAF1}$ ([APC$_{active}$] + [p21$_{active}$]), and their corresponding form of Cyclin B1 (monomer and dimers). Therefore, the terms in Equations (5.1)-(5.5) that represent

degradation (destruction) rate of Cyclin B1 (monomer and dimers) are dependent on the activity of APC and $p21^{CIP1/WAF1}$. That is, the rates of change of the Cyclin B1 monomer and Cyclin B1/CDC2 dimers decrease with the activity of APC and/or $p21^{CIP1/WAF1}$.

The rate at which Cyclin B1 binds to CDC2 (forming the inactive non-phosphorylated complex) is assumed to be proportional, with the association rate k_a, to the product of concentrations of Cyclin B1 (monomer), [CycB], and free CDC2, [CDC2]. In addition, dissociation of Cyclin B1 bound to CDC2 (i.e. [CDC2CycB]) occurs at a first-order rate, with rate constant k_d. According to Pomerening *et al.* [14, 244], the binding affinity of Cyclin B1 to CDC (k_d/k_a) should be less than 1 nM. The constant [$CDC2_{total}$] denotes the total concentration of available CDC2, therefore, the concentration of free CDC2 is:

$$\begin{aligned}[\mathrm{CDC2}](t) = C_2[\mathrm{CDC2_{total}}] - ([\mathrm{CDC2CycB}](t) + [\mathrm{CDC2CycB_{YP}}](t) \\ + [\mathrm{CDC2CycB_{YPTP}}](t) + [\mathrm{CDC2CycB_{YPTP}}](t))\end{aligned} . \qquad (5.11)$$

The parameter k_{Wee1} represents the rate constant for the inactivation of Cyclin B1/CDC2 dimers, namely, the non-phosphorylated complex and the active Cyclin B1/CDC2 complex (T161-phosphorylated complex) by Wee1 kinase. Therefore, the rate of this inactivation is assumed to be proportional, with inactivation rate constant k_{Wee1}, to the product of the concentrations of active Wee1 kinase ([$Wee1_{active}$]) and Cyclin B1/CDC2 complexes ([CDC2CycB] and [$CDC2CycB_{TP}$]). Moreover, the mechanism through which the progress of the cell cycle traverse is switched from S-G_2-M phase-like state to G_1 phase-like state depends on the basal activity of Wee1 kinase. As suggested by Pomerening *et al.* [14], it is assumed that the accumulation of active Cyclin B1/CDC complex and non-phosphorylated Cyclin B1/CDC2 complex (or inactive Wee1 kinase accumulation) triggers the switch from S-G_2-M phase-like state to G_1 phase-like state, because the basal activity of Wee1 kinase is non-zero. Thereby, this accumulation drives the formation of a small amount of inactive Cyclin B1/CDC2 complexes even in the S-G_2-M state. Therefore, given that [$Wee1_{total}$] denotes the total concentration of Wee1 kinase (active and inactive forms), the rates of Wee1 inactivation of active and non-phosphorylated Cyclin B1/CDC2 complexes are assumed to be proportional, with inactivation rate constant $k_{Wee1basal}$, to the product of concentrations of inactive Wee1 kinase ([$Wee1_{total}$] - [$Wee1_{active}$]) and the Cyclin B1/CDC2 complexes ([$CDC2CycB_{TP}$] and [CDC2CycB]).

The parameter k_{CDC25} represents the rate constant for the activation of Cyclin B1/CDC2 dimers, namely, the Y15-phosphorylated complex (inactive complex) and the Y15- and T161-phosphorylated complex (inactive complex) by CDC25 phosphatase. Therefore, the rate of this activation is assumed to be proportional, with inactivation rate constant k_{CDC25}, to the product of the concentrations of active CDC25 phosphatase ([$CDC25_{active}$]) and Cyclin B1/CDC2 complexes ([$CDC2CycB_{YP}$] and [$CDC2CycB_{YPTP}$]). Moreover, the mechanism through which the progress of the cell cycle traverse is switched from G_1 phase-like state to S-G_2-M phase-like state depends on the basal activity of CDC25 phosphatase. Pomerening *et al.* [14] suggested that the accumulation of inactive Cyclin B1/CDC2 complexes (namely, [$CDC2CycB_{YP}$] and [$CDC2CycB_{YPTP}$]) and/or the accumulation of inactive CDC25 phosphatase trigger the switch from G_1 phase-like state to S-G_2-M phase-like state, since the basal activity of CDC25 phosphatase is non-zero. Thereby, this accumulation drives the formation of a small amount of active Cyclin B1/CDC2 complex and non-phosphorylated complex even in G_1 phase. Therefore, given that [$CDC25_{total}$] denotes the total concentration of CDC25 phosphatase (active and inactive forms), the rates of CDC25 activation of inactive Cyclin B1/CDC2 complexes ([$CDC2CycB_{YP}$] and [$CDC2CycB_{YPTP}$]) are assumed to be proportional, with inactivation rate $k_{CDC25basal}$, to the product of concentrations of inactive CDC25 phosphatase ([$CDC25_{total}$] - [$CDC25_{active}$]) and the Cyclin B1/CDC2 complex ([$CDC2CycB_{YP}$] and [$CDC2CycB_{YPTP}$]). The basal activities of Wee1 kinase and CDC25 are given by the following:

$$k_{Wee1basal} = \frac{k_{Wee1}}{factor}$$
$$k_{CDC25basal} = \frac{k_{CDC25}}{factor} \tag{5.12}$$

where the parameter *factor* represents the strength of the feedback from CDC2 to Wee1 and CDC25 which is critical for the bistability of the CDC2/APC system.

In the extended cell cycle model, it is assumed that the dependence of APC, Wee1 kinase and CDC25 phosphatase on CDC2 activity (T161-phosphorylated complex) is described by Hill functions with Hill coefficients denoted by n_{APC}, n_{Wee1} and n_{CDC25}, respectively (see Equations (5.7)-(5.9)) [14]. The EC_{50}'s are the half maximal effective concentrations in the Hill functions, and in the extended model, the EC_{50} concentrations (in nM) in the Hill

functions relating to the rates of change of APC, Wee1 and CDC25 are denoted by $EC50_{APC}$, $EC50_{Wee1}$ and $EC50_{CDC25}$, respectively. The rate constants for turning on the APC, the Wee1 kinase and the CDC25 phosphatase are denoted by k_{APCon}, k_{Wee1on} and $k_{CDC25on}$, respectively. The corresponding rate constants for turning them off (i.e., APC, Wee1 and CDC25) are k_{APCoff}, $k_{Wee1off}$ and $k_{CDC25off}$, respectively.

In the extended cell cycle model of the CDC2/APC system described in this chapter, it is assumed that the dependence of $p21^{CIP1/WAF1}$ on CDC2 activity is described by a Hill function with Hill coefficient denoted by n_{p21} (Equation (5.10)). The EC_{50} concentration (in nM) in the Hill equation relating to the rate of change of p21 is denoted by $EC50_{p21}$. The rate constants for turning the CKI $p21^{CIP1/WAF1}$ on and off are denoted by k_{APCon} and k_{APCoff}, respectively (Equation (5.10)).

Phosphorylation of CDC2 (CDK1) on T161 residue is necessary for CDC2 (CDK1) kinase activity (activation). This site (i.e. T161) is activated (phosphorylated) by CAK in the presence of Cyclin B1 (see Section 3.3) [261-263]. This activation (phosphorylation) process is modelled by a first-order kinetic reaction, where k_{CAK} denotes the rate constant of the phosphorylation of the Y15-phosphorylated complex ($[CDC2CycB_{YP}]$) on T161 residue, resulting in Y15- and T161-phosphorylated complex ($[CDC2CycB_{YPTP}]$). Type 2C protein phosphatases (PP2Cs) are essential physiological regulators of cellular growth and of cellular stress signalling [264]. Additionally, it has been demonstrated that PP2Cs are linked to the cellular stress response (a reaction to the threat of macromolecular damage [265]) by inhibiting the cellular stress signalling. Moreover, it has been suggested that PP2Cs dephosphorylate T-loop-activated kinases (i.e. CDC2 or CDK1) that are involved in DNA damage [266]. This inactivation (dephosphorylation) process is modelled by a first-order kinetic reaction, where k_{PP2C} denotes the rate constant of the dephosphorylation of the Y15- and T161-phosphorylated complex ($[CDC2CycB_{YPTP}]$) on T161, resulting in Y15-phosphorylated complex ($[CDC2CycB_{YP}]$).

The original cell cycle model by Pomerening *et al.* [14] describes the CDC2/APC system of *Xenopus laevis* embryos. The model has been extended to account for the effects of the CKI $p21^{CIP1/WAF1}$ to describe the cell cycle process of human osteosarcoma cells. Although

the key elements and interactions within the cell cycle process are similar in most eukaryotes, the interaction rate constants may vary significantly among organisms [243]. In addition, these rate constants may also vary widely between different cell types of the same organism or within the same cell type under different conditions (e.g. cell size, temperature, pH). Therefore, in the extended mathematical model describing the cell cycle process, all values for the rate constants (which involve time units in them) have been appropriately scaled (re-parameterised) with respect to time with the scaling factor C_1 (see Equations (5.1)-(5.11)), according to the live-cell data provided in Cyclin B1-GFP experiments. The method for calculating the dimensionless constant C_1 is described in Section 5.5. In addition, concentrations of the cell cycle components (i.e. variables) vary drastically among organisms, among different cell types and under different conditions (external and internal). Therefore, in the extended cell cycle model, this has been taken into consideration by appropriately scaling (re-parameterising) all the constant concentrations with the same scaling factor C_2 for simplicity (see Equations (5.1)-(5.11)), according to the live-cell data provided in Cyclin B1-GFP experiments. In the extended model, the scaling (by C_2) includes the total concentrations of CDC2 ($[CDC2_{total}]$), APC ($[APC_{total}]$), Wee1 ($[Wee1_{total}]$), CDC25 ($[CDC25_{total}]$) and $p21^{CIP1/WAF1}$ ($[p21_{total}]$) as well as EC_{50}'s relating to the Hill functions of APC ($[EC50_{APC}]$), Wee1 kinase ($[EC50_{Wee1}]$), CDC25 phosphatase ($[EC50_{CDC25}]$) and CKI $p21^{CIP1/WAF1}$ ($[EC50_{p21}]$).

In the extended cell cycle model, seven new parameters were added to the original cell cycle model by Pomerening *et al.* [14]. Five parameters were added to account for the effects of the CKI $p21^{CIP1/WAF1}$, these parameters are, the Hill coefficient n_{p21}, the half maximal effective concentration $EC50_{p21}$, the total concentration of $p21^{CIP1/WAF1}$ ($[p21_{total}]$), the rate constant for activating $p21^{CIP1/WAF1}$ (k_{p21on}) and the rate constant for switching $p21^{CIP1/WAF1}$ off (k_{p21off}). Two parameters, the scaling factors C_1 (in time) and C_2 (in concentration), were added to the extended cell cycle model to account for the variations among organisms since the extended model presented in this chapter describes the cell cycle process in mammalian cells (in the absence of TPT).

The corresponding initial conditions (i.e. at time t = 0 min) for the *original* cell cycle model in Pomerening *et al.* [14] are given by:

$$\begin{gathered}[\text{CycB}](0)=0,[\text{CDC2CycB}](0)=0,\ [\text{CDC2CycB}_{\text{YP}}](0)=0,\\ [\text{CDC2CycB}_{\text{YPTP}}](0)=0,\ [\text{CDC2CycB}_{\text{TP}}](0)=0,[\text{APC}_{\text{active}}](0)=0,\\ [\text{Wee1}_{\text{active}}](0)=0,\ [\text{CDC25}_{\text{active}}](0)=0 \text{ and } [\text{CDC2}](0)=100\end{gathered} \quad . \tag{5.13}$$

As a first attempt to determine the initial conditions for the extended cell cycle model, it is assumed that $\text{p21}^{\text{CIP1/WAF1}}$ functions in a similar manner to the APC (for simplicity). Accordingly, the initial condition (i.e. at time $t=0$) for the variable relating to the CKI $\text{p21}^{\text{CIP1/WAF1}}$ in the extended model is given by:

$$[\text{p21}_{\text{active}}](0)=0\,. \tag{5.14}$$

Figure 5.12 represents the simulated output of the *extended* model (prior to scaling by C_1 or C_2), using the initial conditions in Equations (5.13) and (5.14) and assuming that the parameters relating to the activity of $\text{p21}^{\text{CIP1/WAF1}}$, namely, n_{p21}, EC50_{p21}, $[\text{p21}_{\text{total}}]$, k_{p21on} and k_{p21off} are equal to their corresponding parameter values relating to the activity of the APC (Table 5.1). Based on the advice of the biological collaborators at Cardiff, the time at which a new cell cycle starts can be identified in the model (Figure 5.12) as the point in time, denoted by t_{G1}, when the concentration of active CDC2/Cyclin B1 (i.e. $[\text{CDC2CycB}_{\text{TP}}]$) is the lowest and non-zero in a complete cell cycle process. In this case, $t_{\text{G1}}=141.75$ min (see Figure 5.12), and accordingly, the values of the variables in Equations (5.1)-(5.11) were evaluated at t_{G1}. Therefore, the corresponding re-parameterised initial conditions (before scaling by C_2) for the *extended* cell cycle model are given by:

$$\begin{gathered}[\text{CycB}](0)=0.141396,[\text{CDC2CycB}](0)=0.992654,\\ [\text{CDC2CycB}_{\text{YP}}](0)=1.4763,[\text{CDC2CycB}_{\text{YPTP}}](0)=11.41023,\\ [\text{CDC2CycB}_{\text{TP}}](0)=1.64121,[\text{CDC2}](0)=84.479606,\\ [\text{APC}_{\text{active}}](0)=3.8539,[\text{Wee1}_{\text{active}}](0)=15,\\ [\text{CDC25}_{\text{active}}](0)=1.54009 \text{ and } [\text{p21}_{\text{active}}](0)=3.8539\end{gathered} \quad . \tag{5.15}$$

This correction (re-parameterisation) step for the initial conditions has been performed because in a real case, a small amount of the active form of Cyclin B1/CDC2 (i.e. $[\text{CDC2CycB}_{\text{TP}}]$) is formed even in the G_1 phase of the cell cycle since the basal activity of CDC25 phosphatase is non-zero. To ensure that this re-parameterisation does not affect the qualitative behaviour of the extended cell cycle model, the output of the model was simulated (prior to scaling by C_1 or C_2) using the initial conditions in Equations (5.15) and assuming that the parameters relating to the activity of $\text{p21}^{\text{CIP1/WAF1}}$, namely, n_{p21}, EC50_{p21},

$[p21_{total}]$, k_{p21on} and k_{p21off} are equal to the corresponding parameters relating to the activity of the APC (Table 5.1). The simulation results of the extended cell cycle model using the re-parameterised initial conditions in Equation (5.15) illustrated in Figure 5.15, reproduce close qualitative agreement to experimentally observed results (i.e. Cyclin B1-GFP experiments in Figure 5.10).

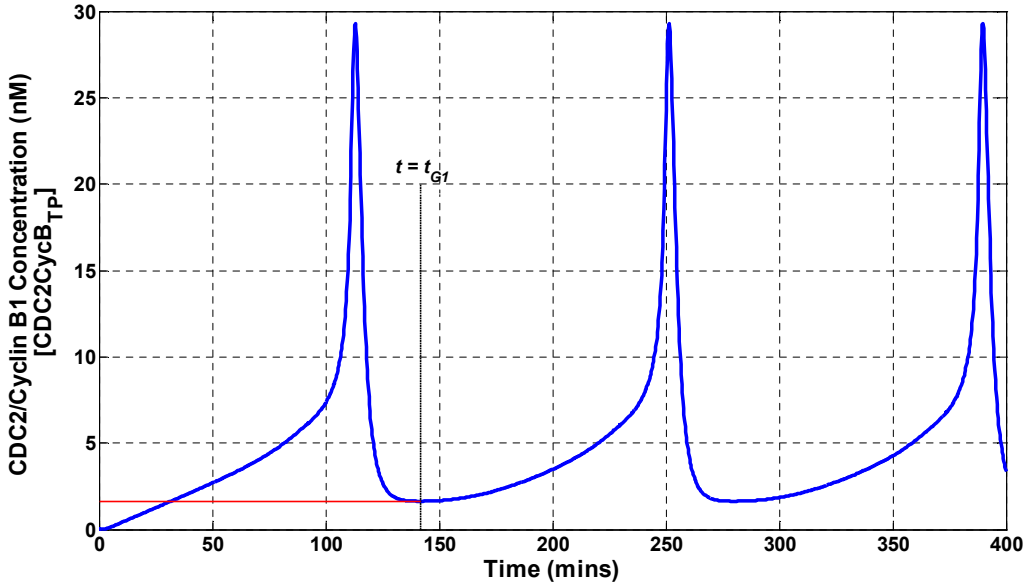


Figure 5.12: Simulation results of the extended cell cycle model to determine the conditions at $t = t_{G1}$. The horizontal sold red line indicates the lowest point in the activity of CDC.

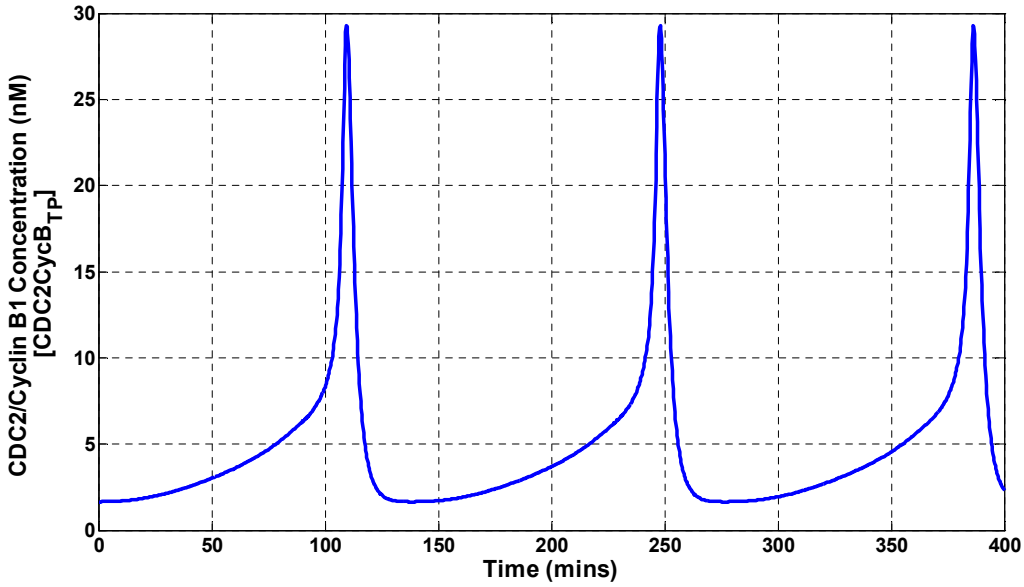


Figure 5.13: Simulation results of the extended cell cycle model using the initial conditions given in Equation (5.14).

In this chapter, parameter estimation was performed for one complete cell cycle starting at G_1 phase of a new born cell, i.e. at time $t = t_{G1}$. Accordingly, in this chapter, time $t = 0$ corresponds to the time when the cell tracking began (at G_2 phase cells), i.e. the time when the first Cyclin B1-GFP fluorescence image was captured (see Section 5.2). This time point, i.e. $t = t_{G1}$, can be determined from the experimental data using the software FLUROTRAK (see Section 5.2), an important feature of this tool is the ability to index any bifurcating point $t = t_B$, where one cell divides into two daughter cells, as shown in Figure 5.13. Therefore, the time at which a new cell cycle starts, $t = t_{G1}$, is the data point after $t = t_B$ (Figure 5.14.), where $t_{G1} - t_B = 20$ min (for each new cycle) since Cyclin B1-GFP fluorescence images were captured every 20 min (see Section 5.2). Moreover, t_{G1} varies among organisms and between different cell types of the same organism or within the same cell type under different conditions. Therefore, the corresponding values for the extended cell cycle model variables (scaled by C_2 to account for variations among organisms and cell types) at the start of a new cell cycle, i.e. at time $t = t_{G1}$ are given by:

$$\begin{aligned}
[\text{CycB}](t_{G1}) &= 0.141396C_2, [\text{CDC2CycB}](t_{G1}) = 0.992654C_2, \\
[\text{CDC2CycB}_{\text{YP}}](t_{G1}) &= 1.4763C_2, [\text{CDC2CycB}_{\text{YPTP}}](t_{G1}) = 11.41023C_2, \\
[\text{CDC2CycB}_{\text{TP}}](t_{G1}) &= 1.64121C_2, [\text{CDC2}](t_{G1}) = 84.479606C_2, \\
[\text{APC}_{\text{active}}](t_{G1}) &= 3.8539C_2, [\text{Wee1}_{\text{active}}](t_{G1}) = 15C_2, \\
[\text{CDC25}_{\text{active}}](t_{G1}) &= 1.54009C_2 \text{ and } [\text{p21}_{\text{active}}](t_{G1}) = 3.8539C_2
\end{aligned} \quad . \qquad (5.16)$$

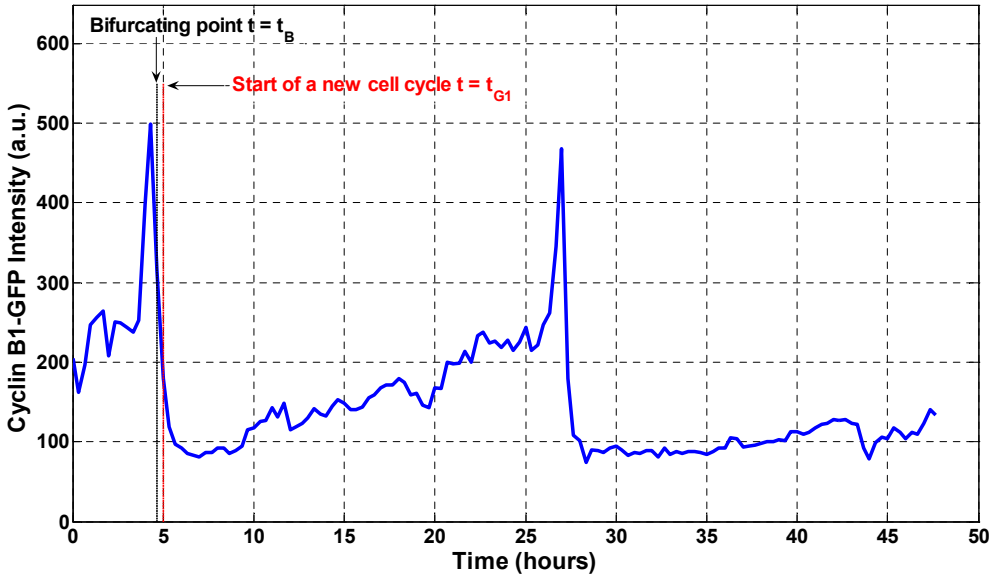


Figure 5.14: Average cytoplasmic Cyclin B1-GFP intensity profile track of the untreated G_2 cell in Figure 5.2, where t_B is the bifurcating point and t_{G1} is the time point at which a new cell cycle starts.

Prior to estimating the unknown model parameters (see Section 5.5), the values for the parameters *factor*, n_{p21} and $EC50_{p21}$ (in the extended cell cycle model) which determine the qualitative features of the output of the extended model were verified by simulating the model output to ascertain the closest qualitative agreement (explosive spikes represented by the G_2/M transition back to G_1) with Cyclin B1-GFP data (see section 5.2). This qualitative computation is insensitive to changes in the values for the unknown model parameters.

The impact of varying the values for the feedback strength (i.e. *factor*) in the extended cell cycle model is shown in Figures 5.15 and 5.16. The parameter *factor* was varied in the range of 1-23. The values are: (a) 1, (b) 5, (c) 7, (d) 10, (e) 15, (f) 20 and (g) 23. If the value for *factor* is 1 (weak feedback) the stimulus/response relationship is monostable and the system reaches a stable steady state. When the level of feedback strength is modestly increased to *factor* = 5, the system oscillates, and the steady-state response is monostable. However, these oscillations are not sustained (i.e., damped oscillations) and the CDC2/APC system reaches a stable steady-state close to the steady state reached by the system when the strength of feedback is 1 (*factor* = 1). Increasing the level of feedback strength to 7, results in a monostable steady-state response but highly ultrasensitive. Moreover, at this level of feedback strength (i.e., *factor* = 7) the system oscillates sustainably and the amplitude of these oscillations increases with the strength of feedback.

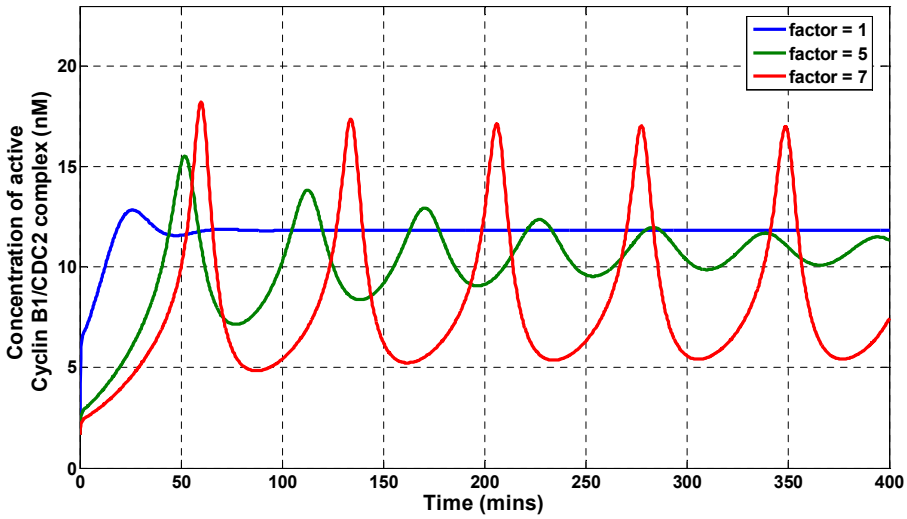


Figure 5.15: Simulation results of Cyclin B1/CDC2 activity in the extended cell cycle model by varying the feedback strength, *factor* = 1 (blue), *factor* = 5 (green) and *factor* = 7 (red).

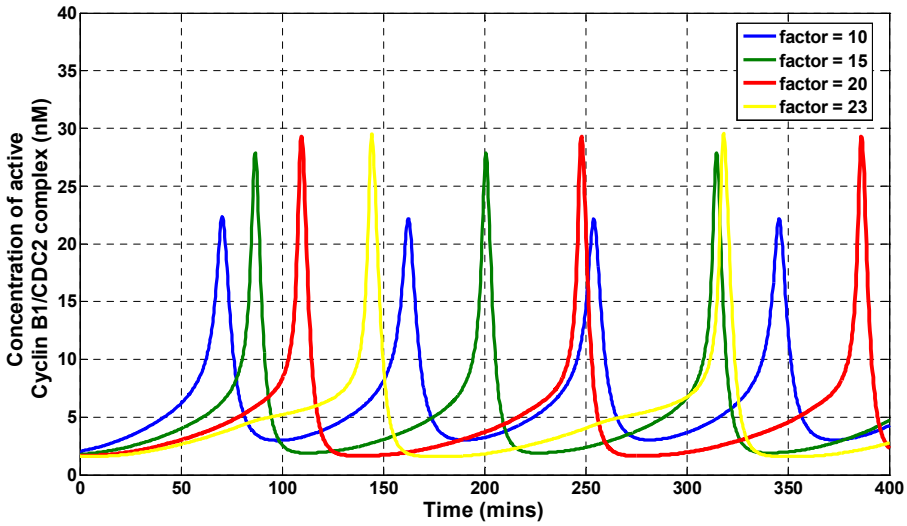


Figure 5.16: Simulation results of Cyclin B1/CDC2 activity in the extended cell cycle model by varying the feedback strength, *factor* = 10 (blue), *factor* = 15 (green), *factor* = 20 (red) and *factor* = 23 (yellow).

Increasing the level of feedback strength (in the extended cell cycle model) to 10, 15, 20 or 23 (*factor* = 10, 15, 20 or 23) results in bistable steady-state responses with explosive spikes of CDC2 activity and more realistic-looking oscillations when compared to Cyclin B1-GFP data (see Figure 5.14). In addition, the amplitudes of these oscillations increase with increasing the level of feedback strength up to the value of 20 (*factor* = 20). Further increase in the level of feedback strength (i.e. when *factor* = 23) results in differences in the amplitudes being indistinguishable. The level of feedback strength also determines the size of the hysteresis loop. That is, for an increase in the feedback strength (*factor*), the hysteresis loop becomes larger [244]. In addition, by increasing the level of feedback strength, the time required before the abrupt G_2/M transition increases. Note that, the exact level of feedback at which the transition from damped oscillations to sustained oscillations depends on the parameter values (including concentrations and rate constants) used in the extended cell cycle model. Setting the level of feedback strength to the value of 20 (*factor* = 20) or 23 (*factor* = 23) results in a closer qualitative agreement to Cyclin B1-GFP data (see Figure 5.14) in comparison to lower levels of feedback strength, particularly in the abrupt G_2/M transition back to G_1 phase. However, when the value of the parameter *factor* = 20 the slope of the interphase (i.e., G_1-S-G_2) is steeper than the slope of the interphase when the value of *factor* = 23. The steepness of the interphase slope when the value of level of feedback strength is 20 (i.e., *factor* = 20) reproduces closer qualitative agreement to Cyclin B1-GFP data. Therefore, the level of feedback strength used in the extended cell cycle model in this thesis is 20 (i.e., the value of the parameter *factor* = 20).

The impact of varying the values for the Hill coefficient n_{p21} is shown in Figures (5.17). The parameter n_{p21} was varied in the range of 1-6. The values are: (a) 1, (b) 2, (c) 3, (d) 4, (e) 5 and (f) 6. If the value for $n_{p21} = 1$ and 2, the stimulus/response relationship is monostable (with no oscillations) and the system reaches a stable steady-state. When the value of n_{p21} is increased to 3 ($n_{p21} = 3$), the system oscillates, and the steady-state response is monostable. However, these oscillations are not sustained (i.e., damped oscillations) and the CDC2/APC system reaches a stable steady-state. Increasing the Hill coefficient to 4, 5 or 6 (n_{p21} = 4, 5 or 6) results in bistable steady-state responses with explosive spikes of CDC2 activity and more realistic-looking oscillations when compared to Cyclin B1-GFP data (see Figure 5.14). Moreover, the amplitudes of these oscillations increase with increasing value of n_{p21} up to n_{p21} = 5. That is, any increase in the Hill coefficient n_{p21} (beyond $n_{p21} = 5$) has a relatively indistinguishable affect on the dynamical response of the cell cycle model in comparison to the response of the system when $n_{p21} = 5$. Therefore empirically, the value of the parameter n_{p21} chosen in the cell cycle presented in this thesis is 5 (i.e., $n_{p21} = 5$).

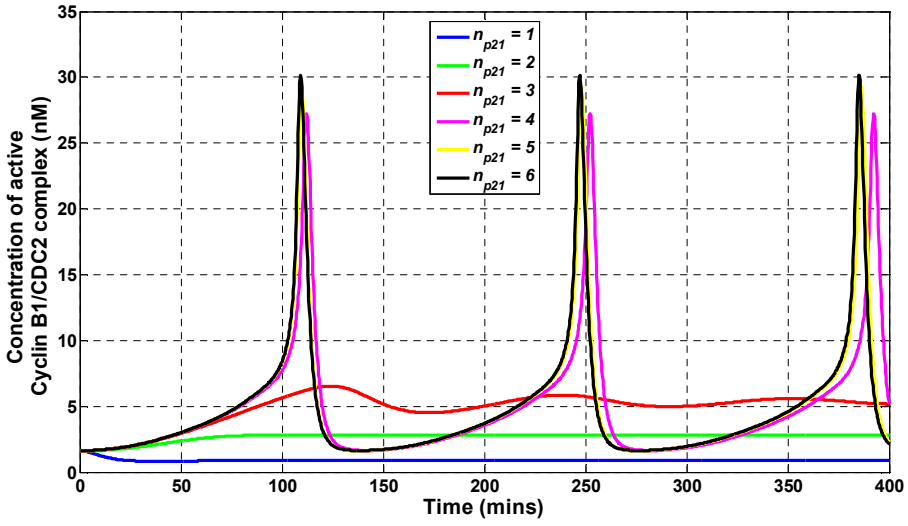


Figure 5.17: Simulation results of Cyclin B1/CDC2 activity by varying the Hill coefficient n_{p21}. These values are, $n_{p21} = 1$ (blue), $n_{p21} = 2$ (green), $n_{p21} = 3$ (red), $n_{p21} = 4$ (magenta), $n_{p21} = 5$ (yellow) and $n_{p21} = 6$ (black).

The impact of varying the values for the half maximal effective concentration $EC50_{p21}$ is shown in Figures (5.18). The parameter $EC50_{p21}$ was varied in the range of 10 nM-50 nM. The values are: (a) 10 nM, (b) 20 nM, (c) 30 nM, (d) 40 nM and (e) 50 nM. If the value for $EC50_{p21}$ = 10 nM, the stimulus/response relationship is monostable (with no oscillations)

and the system reaches a stable steady-state. Increasing the half maximal effective concentration $EC50_{p21}$ to 20 nM, 30 nM, 40 nM and 50 nM ($EC50_{p21}$ = 20 nM, 30 nM, 40 nM or 50 nM) results in bistable steady-state responses with explosive spikes of CDC2 activity and more realistic-looking oscillations when compared to Cyclin B1-GFP data (see Figure 5.14). Moreover, the amplitudes of these oscillations (dynamical response) increase with increasing the value of the half maximal effective concentration parameter $EC50_{p21}$. In comparison between the simulation results (dynamical responses) of the CDC2/APC system shown in Figure 5.18 for the different values of $EC50_{p21}$, namely, $EC50_{p21}$ = 30 nM, 40 nM and 50 nM, the steep transition (slope) from S-G_2-M phase to G_1 phase using the value of $EC50_{p21}$ = 30 nM results in a closer qualitative agreement with the transition from S-G_2-M phase to G_1 in the Cyclin B1-GFP data (Figure 5.14). Therefore, the value used for the half maximal effective concentration $EC50_{p21}$ in the extended cell cycle model is 30 nM (i.e., $EC50_{p21}$ = 30 nM).

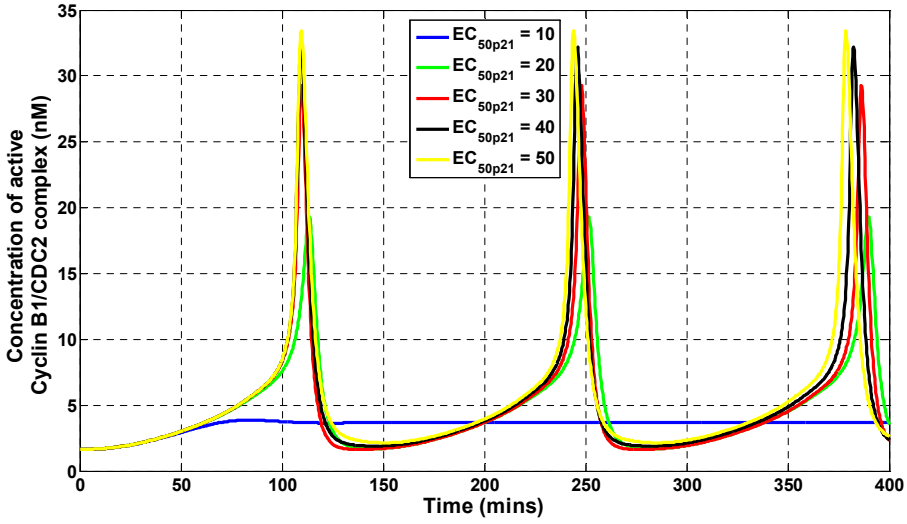


Figure 5.18: Simulation results of Cyclin B1/CDC2 activity by varying the half maximal effective concentration $EC50_{p21}$. These values are, $EC50_{p21}$ = 10 (blue), $EC50_{p21}$ = 20 (green), $EC50_{p21}$ = 30 (red), $EC50_{p21}$ = 40 (black) and $EC50_{p21}$ = 50 (yellow).

In the extended model, the rate constant for activating (turning on) $p21^{CIP1/WAF1}$ (k_{p21on}) is assumed to be equal to the rate constants for activating the APC, Wee1 kinase and CDC25 phosphatase. Therefore, the value used for k_{p21on} is 0.8 min^{-1}. The total concentration of $p21^{CIP1/WAF1}$ ($[p21_{total}]$), and the rate constant for switching $p21^{CIP1/WAF1}$ off (k_{p21off}) are unknown parameters in the model which were estimated using the time-lapse microscopy data (Cyclin B1-GFP fluorescence data). Table 5.1 represents the known parameter values

of the extended cell cycle model for live human osteosarcoma cells: values for the known parameters are characterised by those given in Pomerening *et al.* [14] in addition to the parameters evaluated for the extended cell cycle model, namely *factor*, n_{p21}, $EC50_{p21}$, and k_{p21on}. All extended cell cycle model parameters (known and unknown) have been implicitly scaled by C_1 or C_2 in Equations (5.1)-(5.12) and Equation (5.16).

Table 5.1: Table of known parameter values of the extended cell cycle model for live human osteosarcoma cells.

Parameter	Value
k_{synth} (nM min^{-1})	1.2
k_{dest} (nM^{-1} min^{-1})	0.005
k_a (nM^{-1} min^{-1})	0.1
k_d (min^{-1})	0.001
Factor	20
k_{Wee1} (nM^{-1} min^{-1})	0.1
k_{CDC25} (nM^{-1} min^{-1})	0.1
$CDC2_{total}$ (nM)	100
APC_{total} (nM)	50
$Wee1_{total}$ (nM)	15
$CDC25_{total}$ (nM)	15
n_{APC}	5
n_{Wee1}	4
n_{CDC25}	4
n_{p21}	5
$EC50_{APC}$ (nM)	30
$EC50_{Wee1}$ (nM)	25
$EC50_{CDC25}$ (nM)	25
$EC50_{p21}$ (nM)	30
k_{APCon} (min^{-1})	0.8
k_{Wee1on} (min^{-1})	0.8
$k_{CDC25on}$ (min^{-1})	0.8
k_{p21on} (min^{-1})	0.8
k_{APCoff} (min^{-1})	0.08
$k_{Wee1off}$ (min^{-1})	0.08
$k_{CDC25off}$ (min^{-1})	0.08
k_{CAK} (min^{-1})	0.8
k_{PP2C} (min^{-1})	0.008

The extended cell cycle model of the CDC2/APC system is similar in many ways to the model proposed by Novak and Tyson [221] and refined by Marlovits *et al.* [267]. However, the main difference between the Novak and Tyson model [221] and the model presented in this thesis is in the assumed source of ultra-sensitivity [14]. Moreover, the extended CDC2/APC model is also similar to the Tyson and Novak model [157] in that both rely on a hysteresis mechanism and the governing equations account for the regulation of CDK1/Cyclin B1.

Figure 5.19 illustrates a comparison between the simulation results of the extended cell cycle model of the CDC2/APC system and the original cell cycle model of the CDC2/APC system model by Pomerening *et al.* [14], before scaling by C_1 or C_2. The cell cycle models (extended and original) were simulated using parameter values in Table 5.1 and the re-parameterised initial conditions in Equation (5.16). In simulating the extended cell cycle model, it is assumed that the values for the unknown model parameters $p21_{total}$ and k_{p21off} are equal to their corresponding parameter values relating to the activity of the APC (see Table 5.1)

Indeed, the inhibition effect of the CKI $p21^{CIP1/WAF1}$ plays a significant role in stabilising the G_1 state. The principal effect of adding the negative-feedback loop (inhibition effect of $p21^{CIP1/WAF1}$) is an extension between the two mitotic landmarks, and an increase in IMT, including an extended delay in G_1 state. In the extended cell cycle model, the inter-mitotic time (IMT), denoted by IMT_2 is equal to 138.5 min (i.e. $IMT_2 = 138.5$ min) and in the original cell cycle model, the IMT, denoted by IMT_1 is equal to 131.25 min (i.e., $IMT_1 = 131.25$ min), this result is consistent with the predicted response to the inhibitor $p21^{CIP1/WAF1}$. In addition, the extended cell cycle model reproduces closer qualitative agreement to experimentally observed results than the original cell cycle model (see Figure 5.14) particularly in the abrupt G_2/M transition back to G_1 phase (switch-like response).

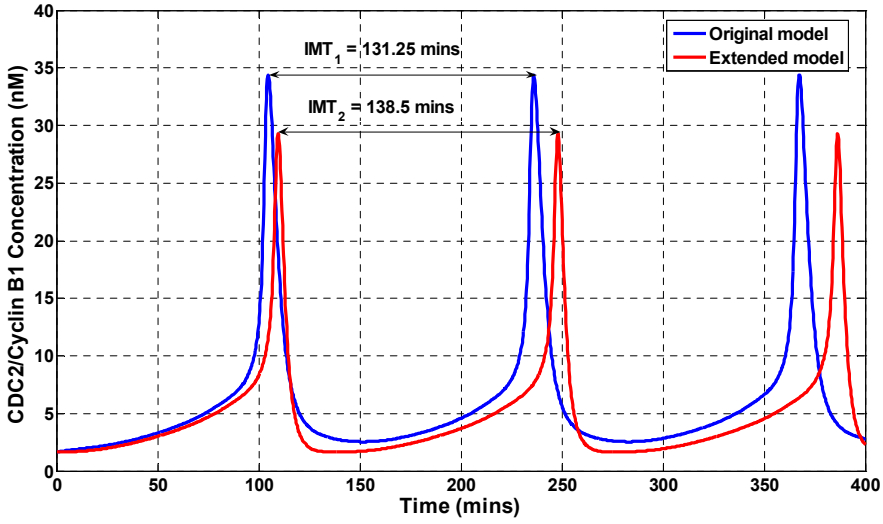


Figure 5.19: Simulation results of extended cell cycle model (in red) and the original cell cycle model (in blue) by Pomerening *et al.* [14].

In all of the time-lapse microscopy experiments (from U-2 OS cell) considered in this chapter, fluorescence intensities (which reflect the concentrations of the active form of Cyclin B1, i.e., [$CDC2CycB_{TP}$]) are measured in arbitrary units (a.u.) and time in minutes (min). The thirty two parameters considered in the extended cell cycle model are:

- k_{synth} and k_{dest}, units: nM min^{-1} (nM per minute) and $(nM\ min)^{-1}$, respectively. Rate constants for Cyclin B1 (monomer) synthesis and Cyclin B1 (monomer and dimers) destruction, respectively;
- k_a and k_d, units: $(nM\ min)^{-1}$ and min^{-1}, respectively. Rate constants for association of CDC2 (monomer) to Cyclin B1 (monomer) forming the inactive non-phosphorylated complex (CDC2CycB) and the dissociation of this complex to the monomers (i.e., Cyclin B1 and CDC2), respectively;
- *factor*, dimensionless constant. This parameter represents the strength of the feedback from CDC2 to Wee1 kinase and CDC25 phosphatase;
- k_{Wee1}, units: $(nM\ min)^{-1}$. Rate constant for the inactivation of Cyclin B1/CDC2 dimers, namely, the non-phosphorylated complex (CDC2CycB) and the active Cyclin B1/CDC2 complex ($CDC2CycB_{TP}$) by Wee1 kinase;
- k_{CDC25}, units: $(nM\ min)^{-1}$. Rate constant for the activation of Cyclin B1/CDC2 dimers, namely, the Y15-phosphorylated complex ($CDC2CycB_{YP}$) and the Y15- and T161-phosphorylated complex ($CDC2CycB_{YPTP}$) by CDC25 phosphatase;
- $CDC2_{total}$, APC_{total}, $Wee1_{total}$, $CDC25_{total}$ and $p21_{total}$, units: nM. The total concentrations of CDC2, APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- n_{APC}, n_{Wee1}, n_{CDC25}, and n_{p21}, dimensionless constants. These dimensionless constants are the Hill coefficients of the Hill functions describing the rates of change of the active forms of the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- $EC50_{APC}$, $EC50_{Wee1}$, $EC50_{CDC25}$, and $EC50_{p21}$, units: nM. The half maximal effective concentrations (EC_{50}'s) in the Hill functions describing the rates of change of the active forms of the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;

- k_{APCon}, k_{Wee1on}, $k_{CDC25on}$ and k_{p21on}, units: min^{-1}. Rate constants for turning on the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- k_{APCoff}, $k_{Wee1off}$, $k_{CDC25off}$ and k_{p21off}, units: min^{-1}. Rate constants for turning off the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- k_{CAK} and k_{PP2C}, units: min^{-1}. The rate constants for phosphorylation/ dephosphorylation of the Y15-phosphorylated complex ($CDC2CycB_{YP}$)/ Y15- and T161-phosphorylated complex ($CDC2CycB_{YPTP}$), respectively.
- C_1 and C_2, dimensionless constants. Scaling factors to account for the variations in the interaction rate constants and the variations in the concentrations of the cell cycle components (among different organisms), respectively.

Three of these parameters are unknown (C_2, k_{p21off} and $p21_{total}$) in addition to two observational parameters (α and β, see Section 5.4 and Section 5.5) are unknown, and need to be estimated from experimental data. The time-lapse microscopy experiments allow direct measurement of the fluorescence intensities (Cyclin B1-GFP) of concentrations of the active form of Cyclin B1 (Cyclin B1/CDC2) in the nucleus and cytoplasm (see Section 5.2). Prior to the parameter estimation for the extended cell cycle model, a structural identifiability analysis has been performed (Section 5.4) to test whether the unknown model parameters are uniquely determined by the observations structure corresponding to the time-lapse microscopy experiments. In the following section, the Taylor series approach is used to show that the extended cell cycle model (for a single cell) is structurally globally identifiable.

5.4 Structural identifiability analysis of the cell cycle model

In this section, the Taylor series approach (see Section 2.4) is applied to the extended cell cycle model at time $t = t_{G1}$ (start of a new cell cycle, see Figure 5.14) described by Equations (5.1)-(5.11). For this study, the output function, $y(t, \mathbf{p})$, of the extended model is taken to be a linear function of active Cyclin B1 concentration ($[CDC2CycB_{TP}]$), that is:

$$y(t, \mathbf{p}) = \alpha CDC2CycB_{TP}(t, \mathbf{p}) + \beta \tag{5.17}$$

where, α and β are the observational parameters estimated from the experimental data. The state-space model describing cell cycle process for a mammalian cell (in the absence of TPT) in this chapter is an uncontrolled (referred to as free or autonomous) non-linear system.

The vector **p** comprising the five unknown parameters of the model is given by:

$$\mathbf{p} = (\alpha, \beta, C_2, k_{\text{p21off}}, \text{p21}_{\text{total}})^{\mathsf{T}}.$$

The unknown model parameters represent observational parameters, a scaling factor, a rate constant and a concentration, and are therefore positive. This means that the set of feasible parameter vectors, Ω, comprises the vectors $(p_1, \ldots, p_5)^{\mathsf{T}}$ such that $p_i > 0$ $(1 \leq i \leq 5)$. The state space vector $\mathbf{x}(t, \mathbf{p})$ is given by:

$$\begin{aligned}\mathbf{x}(t,\ \mathbf{p}) = (&\text{CycB}(t,\ \mathbf{p}), \text{CDC2CycB}(t,\ \mathbf{p}), \text{CDC2CycB}_{\text{YP}}(t,\ \mathbf{p}),\\ &\text{CDC2CycB}_{\text{YPTP}}(t,\ \mathbf{p}), \text{CDC2CycB}_{\text{TP}}(t,\ \mathbf{p}), \text{CDC2}(t,\ \mathbf{p}),\\ &\text{APC}_{\text{active}}(t,\ \mathbf{p}), \text{Wee1}_{\text{active}}(t,\ \mathbf{p}), \text{CDC25}_{\text{active}}(t,\ \mathbf{p}), \text{p21}_{\text{active}}(t,\ \mathbf{p}))^{\mathsf{T}}\end{aligned} \tag{5.18}$$

and the set $M(\mathbf{p})$ is $\mathbb{R}^{10}$. The values for the state variables of the model at $t = t_{\text{G1}}$ are given by:

$$\begin{aligned}\mathbf{x}(t_{\text{G1}},\ \mathbf{p}) = (&0.141396C_2, 0.992654C_2, 1.4763C_2, 11.41023C_2, 1.64121C_2,\\ &84.479606C_2, 3.8539C_2, 15C_2, 1.54009C_2, 3.8539C_2)^{\mathsf{T}}.\end{aligned} \tag{5.19}$$

In this model (i.e., the extended cell cycle model), the values for the state variables at time $t = t_{\text{G1}}$ are actually dependent on components of the vector of unknown parameters **p**.

An arbitrary parameter vector is denoted by $\bar{\mathbf{p}}$ such that:

$$\bar{\mathbf{p}} = (\bar{\alpha}, \bar{\beta}, \bar{C}_2, \bar{k}_{\text{p21off}}, \bar{\text{p}}21_{\text{total}})^{\mathsf{T}}$$

for which, $\text{y}(t,\ \mathbf{p}) = \text{y}(t,\ \bar{\mathbf{p}})$ for all $t \geq 0$. Similar to the structural identifiability analysis for the cell-based kinetics model in Chapter 4, the computer algebra system MATHEMATICA [206] was used to perform the following structural identifiability analysis of the extended cell cycle model (see Appendix A). As explained in Section 2.4, uniqueness of the coefficients in the Taylor series expansions of $y(t, \mathbf{p})$ entails the following, if $\bar{\mathbf{p}} \in \boldsymbol{\Omega}$ is such that $\text{y}(t, \mathbf{p}) = \text{y}(t,\ \bar{\mathbf{p}})$ for all $t \geq 0$, then, for each $k = 1, 2, 3, \ldots$

$$y^{(k)}(t_{G1}, \mathbf{p}) = y^{(k)}(t_{G1}, \bar{\mathbf{p}}) \text{ where, } y^{(0)}(t_{G1}, \mathbf{p}) = y(t_{G1}, \mathbf{p}). \tag{5.20}$$

Setting $k = 0$ in Equation (5.20) yields the first Taylor series coefficient given by:

$$y^{(0)}(t_{G1}, \mathbf{p}) = 1.64121 C_2 \alpha + \beta \tag{5.21}$$

and substituting in Equation (5.20) gives:

$$\bar{\beta} = 1.64121 C_2 \alpha - 1.64121 \bar{C}_2 \bar{\alpha} + \beta. \tag{5.22}$$

The 1st derivative term ($k = 1$) gives:

$$y^{(1)}(t_{G1}, \mathbf{p}) = 2.52518 C_2^{\,2} \alpha \tag{5.23}$$

and using the relations in Equation (5.22) yields:

$$\bar{\alpha} = \frac{C_2^{\,2} \alpha}{\bar{C}_2^{\,2}}. \tag{5.24}$$

Considering the relations (Equations (5.22) and (5.24)) between parameters in $\mathbf{p}$ and $\bar{\mathbf{p}}$, the following equations must hold to satisfy Equation (5.20) for $k = 0, 1$:

$$\bar{\alpha} = \frac{C_2^{\,2} \alpha}{\bar{C}_2^{\,2}}, \text{ and } \bar{\beta} = 1.64121 C_2 \alpha - 1.64121 \frac{C_2^{\,2} \alpha}{\bar{C}_2} + \beta. \tag{5.25}$$

The 2nd derivative term ($k = 2$) gives:

$$y^{(2)}(t_{G1}, \mathbf{p}) = (0.107836 - 0.388901 C_2) C_2^{\,2} \alpha \tag{5.26}$$

and using the relations in Equations (5.25) and (5.26) yields:

$$\bar{C}_2 = C_2. \tag{5.27}$$

Considering the relations (Equations (5.25) and (5.27)) between parameters in $\mathbf{p}$ and $\bar{\mathbf{p}}$, the following equations must hold to satisfy Equation (5.20) for $k = 0, 1, 2$:

$$\bar{C}_2 = C_2, \ \bar{\alpha} = \alpha, \text{ and } \bar{\beta} = \beta. \tag{5.28}$$

The 3rd derivative term ($k = 3$) gives:

$$y^{(3)}(t_{G1}, \mathbf{p}) = 0.0598942(-0.0627616 + C_2) C_2^{\,2} (4.45049 + C_2) \alpha \tag{5.29}$$

and substituting in Equation (5.20) yields no information since none of the remaining unknown model parameters are present in the 4th coefficient of the Taylor series expansion. Equating the 4th derivative terms ($k = 4$) gives:

$$y^{(4)}(t_{G1},\mathbf{p}) = 0.00252426C_2^{\,2}\alpha - 0.0345519C_2^{\,3}\alpha - 0.171215C_2^{\,4}\alpha \\ -0.00922425C_2^{\,5}\alpha \tag{5.30}$$

and substituting in Equation (5.20) yields no information since none of the remaining unknown model parameters are present in the 5th coefficient of Taylor series expansion. Equating the 5th derivative terms ($k = 5$) gives:

$$y^{(5)}(t_{G1},\mathbf{p}) = C_2^{\,2}\alpha(-0.000310113 + 2.6964C_2 - 2.50228C_2^{\,3} \\ +0.00142062C_2^{\,4} + C_2^{\,2}(0.0542277 - 5.16716\times10^{-10}\,\mathrm{p21}_{\mathrm{total}})) \tag{5.31}$$

and using the relations in Equations (5.28) yields:

$$\bar{k}_{\mathrm{p21off}} = k_{\mathrm{p21off}} - 1.01719\times10^{-7}\,\mathrm{p21}_{\mathrm{total}} + 1.01719\times10^{-7}\,\bar{\mathrm{p}}21_{\mathrm{total}}. \tag{5.32}$$

Finally, equating the seventh set of coefficients ($k = 6$), and using the relations in Equations (5.28) and (5.32), yields:

$$\bar{\mathrm{p}}21_{\mathrm{total}} = \mathrm{p21}_{\mathrm{total}}. \tag{5.33}$$

Therefore, the following relations must hold in order to satisfy Equation (5.20) for all $k = 0, \ldots, 6$:

$$\bar{C}_2 = C_2,\ \bar{k}_{\mathrm{p21off}} = k_{\mathrm{p21off}},\ \bar{\mathrm{p}}21_{\mathrm{total}} = \mathrm{p21}_{\mathrm{total}},\ \bar{\alpha} = \alpha,\ \text{and}\ \bar{\beta} = \beta. \tag{5.34}$$

Equation (5.20) is therefore satisfied for all $k \geq 0$ with $\mathbf{p} = \bar{\mathbf{p}}$. This is true for generic $\mathbf{p} \in \Omega$, thus the model is structurally globally identifiable, that is all of the model parameters are uniquely determined by the output structure corresponding to the proposed experiments to collect data for the purpose of parameter estimation. In the next section, the unknown model parameters are estimated from time-lapse microscopy data (Cyclin B1-GFP data). The structural identifiability analysis was performed in this section to approach the estimation of the unknown model parameters with greater confidence.

5.5 Parameter estimation of the cell cycle model

In this chapter, the cell cycle model describes the response of growth of single human osteosarcoma cells in the absence of the anti-cancer agent TPT. The unknown model

parameters (as in Section 5.4) were estimated from the Cyclin B1-GFP data of a control cell (untreated cell) for one complete cell cycle event of a newborn cell starting at time $t = t_{G1}$ (see Figures 5.14 and 5.20) to investigate the effects of the CKI $p21^{CIP1/WAF1}$ on the progression of the cell cycle traverse. The expression of Cyclin B1 is driven by a cell-cycle-phase-specific promoter that initiates expression at the end of the S phase of the cell cycle process and peaks during the G_2 phase [268]. Once Cyclin B1 is expressed, it shuttles between the nucleus and cytoplasm during the G_2 phase of the cell cycle traverse. However, Cyclin B1 localisation is primarily cytoplasmic because the rate of nuclear export is much greater than its import [159]. Thereafter (i.e., at the start of the M phase), Cyclin B1 accumulates, in a phosphorylation dependent manner [269, 270]. Accordingly, the G_2/M transition is marked by the localisation of Cyclin B1 in the cell. Since the unknown model parameters are estimated between two mitotic events, Cyclin B1-GFP data in the cytoplasm were averaged (i.e., Cyclin B1-GFP data from the two cytoplasmic ROI's, see Section 5.2) and used for the curve fitting as shown in Figure 5.20.

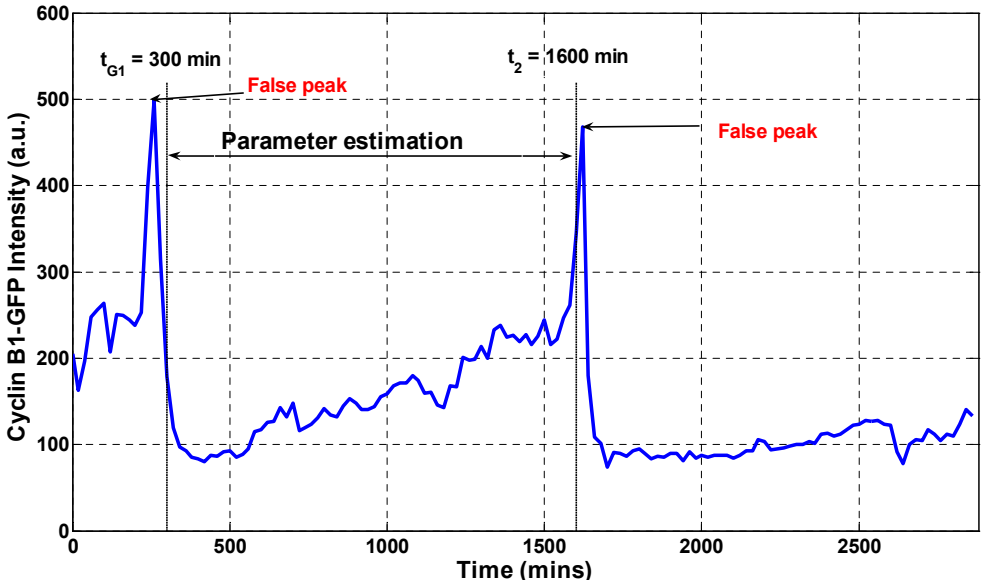


Figure 5.20: Average cytoplasmic Cyclin B1-GFP fluorescence intensities for an untreated cell.

The spikes in the Cyclin B1-GFP experimental data correspond to morphological changes in the cell during the M phase of the cell cycle traverse. As cells grow, their shapes change and round up just before cell division (i.e., bifurcation or cytokinesis) and the spikes in the Cyclin B1-GFP (at mitosis) are artefacts due to these morphological changes (rounding up of cells). This is because the light reflected back (measure of the fluorescence) is at its

maximum intensity at this stage of the cell cycle process and the camera takes full thickness image, and so for every pixel the volume of cytoplasm increases resulting in a false peak (see Figure 5.20) in the Cyclin B1-GFP data. The location of the false peak (in time) can be determined from the experimental data using the software FLUROTRAK. Therefore, the unknown model parameters were estimated between the two vertical dashed lines in Figure 5.20 staring from the G_1 phase (at time $t = t_{G1}$) of the cell cycle (newborn daughter cell) up to the experimental point in time before the false peak, denoted by t_2 (see Figure 5.20).

Parameter estimation was performed using the commercial simulation software package FACSIMILE (MCPA Software, U.K.). The optimisation method used to obtain parameter estimates involves the minimisation of the weighted least-squares criterion, see Section 2.5. The unknown parameters of the extended cell cycle model, namely, the observational parameters (α and β), the concentration scaling factor C_2 (specific to cell type and cellular conditions), total concentration of the CKI $p21^{CIP1/WAF1}$ ($[p21_{total}]$) and the rate constant for turning $p21^{CIP1/WAF1}$ off (k_{p21off}) are estimated from the experimental data (see Section 5.2). Live cells are tracked in time and therefore, these parameters (i.e., the unknown model parameters) are permitted to change (if appropriate) at cell division (cytokinesis), however, they remain constant during an individual cycle. The parameter values in Table 5.1 are specific for all cell types and cellular conditions (internal and external) that influence the cell cycle progression, in addition, these parameters (in Table 5.1) are also permitted to change implicitly (if appropriate) at cell division since they are scaled by C_1 or C_2. However, they remain constant during an individual cell cycle.

For the Cyclin B1-GFP data of the control (untreated) cell shown in Figure 5.20, the progenitor cell divides into two daughter cells at 280 min (i.e., t_B = 280 min, see Figure 5.14) after the start of the experiment, and therefore, in this particular track, t_{G1} = 300 min (see Figure 5.20). The false peak is located at time t = 1620 min and therefore, t_2 = 1600 min (the experimental data point just before the false peak), since the fluorescence image sequences were captured every 20 min (see Section 5.2 and Figures 5.14 and 5.20). In addition, the untreated cell divides again at time t = 1640 min (second cell division). Therefore, in this particular fitting, the unknown model parameters were estimated between t_{G1} = 300 min and t_2 = 1600 min. Then, the estimated parameters were used to

simulate the model from t_2 = 1600 min to t_3 = 1660 min (which is t_{G1} of the new cycle) to ensure that the cell has completed the cycle (i.e., the cell has completely divided into two daughter cells). The scaling factor C_1 is given by:

$$C_1 = \frac{\mathrm{IMT}_a}{\mathrm{IMT}_b} \tag{5.35}$$

where, IMT_a is the inter-mitotic time in the extended model before scaling and IMT_b is the inter-mitotic time in the experimental data (i.e., the time for one individual cycle). IMT_a is a constant and is equal to 138.5 min (see Figure 5.19) and in this particular track (Figure 5.20), IMT_b = 1360 min (1640 min – 280 min), and therefore, C_1 = 0.1017 (Table 5.2).

Table 5.2: Best parameter estimates for control (untreated) cell data obtained for the model Equations (5.1)-(5.11) and state values at $t = t_{G1}$ in Equation (5.16), estimated using live-cell Cyclin B1-GFP data.

Parameter				**Value**
C_1				0.1017
Parameter	**Value**	**SDLN**	**5%**	**95%**
α (a.u. nM^{-1})	$2.2041\times10^{+1}$	0.0948	$1.8860\times10^{+1}$	$2.5759\times10^{+1}$
C_2	9.2038×10^{-1}	0.0179	8.9363×10^{-1}	9.4793×10^{-1}
β (a.u.)	$6.7686\times10^{+1}$	0.0925	$5.8132\times10^{+1}$	$7.8809\times10^{+1}$
k_{p21off} (min^{-1})	5.2973×10^{-2}	0.5973	1.9829×10^{-2}	1.4152×10^{-1}
$\mathrm{p21}_{\mathrm{total}}$ (nM)	$5.2179\times10^{+2}$	0.4153	$2.6352\times10^{+2}$	$1.0332\times10^{+3}$
RSS = $1.5804\times10^{+1}$				

The fitted (estimated) parameter values, with estimates for their confidence levels are presented in Table 5.2, where it is seen that all parameters are well-determined by the live-cell data. Additionally, three of the estimated parameters, namely C_2, α and β, have low SDLN values corresponding to high confidence in their values. However, the estimated parameter values for k_{p21off} and $\mathrm{p21}_{\mathrm{total}}$ have high SDLN values due to the low signal-to-noise ratio in the experimental data. Moreover, the value of $\mathrm{CORI}_{\mathrm{mean}}$ = 1.09241 indicates that the residuals are acceptable and evenly distributed above and below the (simulated) output of the model and therefore, the resulting fit is good.

The plots shown in Figure 5.21 show a typical trace for the growth of a cell in the absence of the anti-cancer drug TPT. The live-cell data are denoted by the blue circles and the simulated data by the solid red curve. The model fit in Figure 5.21 shows that there is close reproduction of the experimental data by simulated output from the extended cell cycle

model with parameters taking values from Tables 5.1 and 5.2. In an attempt to minimise the impact of the artefact of morphological changes (rounding up) during the final stages of the M-phase of the cell cycle traverse which results in the over-expression of Cyclin B1-GFP fluorescence intensities (false spikes) in the experimental data, the false peak (spike) in the data is under-predicted (minimised or capped) by the extended model.

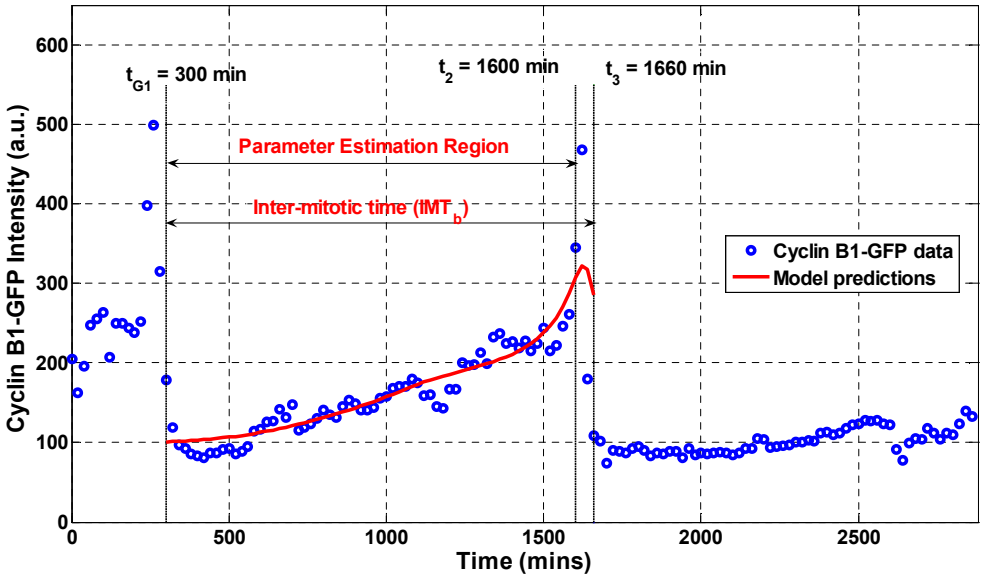


Figure 5.21: Model fit showing model prediction (solid red line) and Cyclin B1-GFP data (blue circles) in the absence of TPT. An artefact of morphological changes during the final stages of mitosis is the over-expression of Cyclin B1 fluorescence giving rise to false peaks in the experimental data.

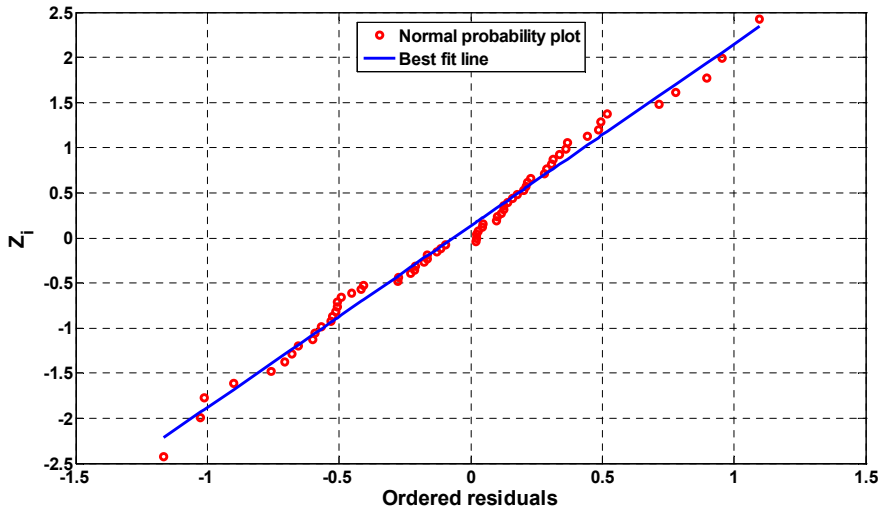


Figure 5.22: A normal probability plot for the residuals in the weighted non-linear least-squares fit of the extended cell cycle model to Cyclin B1-GFP data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

Figure (5.22) is a normal probability plot for the residuals in the weighted non-linear least-squares fit of extended cell cycle model to Cyclin B1-GFP data. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale. The figure suggests that the residuals are normally distributed (with approximately zero mean) since the graph is approximately linear. This is an indicator for the appropriateness of the standard error σ used for each observation y.

The estimated correlation matrix for the well-determined parameters is given in Table 5.3. The parameter pairing that exhibits noticeable (highest) correlation is $p21_{total}$ and k_{p21off} (0.881). These two parameters (i.e., $p21_{total}$ and k_{p21off}) represent the total concentration of CKI $p21^{CIP1/WAF1}$ and the rate at which this CKI is turned off. The high correlation of the logarithm of these parameters indicates that it is the product of these that is important for the fitting, that is, the rate at which the inhibitor is activated.

Table 5.3: The estimated correlation matrix for the well-determined parameters for an untreated (control) cell.

	A	C_2	β	k_{p21off}	$p21_{total}$
α	1	-0.108	-0.687	-0.544	-0.617
C_2	-0.108	1	0.44	-0.579	-0.156
β	-0.687	0.44	1	-0.135	-0.019
k_{p21off}	-0.544	-0.579	-0.135	1	0.881
$p21_{total}$	-0.617	-0.156	-0.019	0.881	1

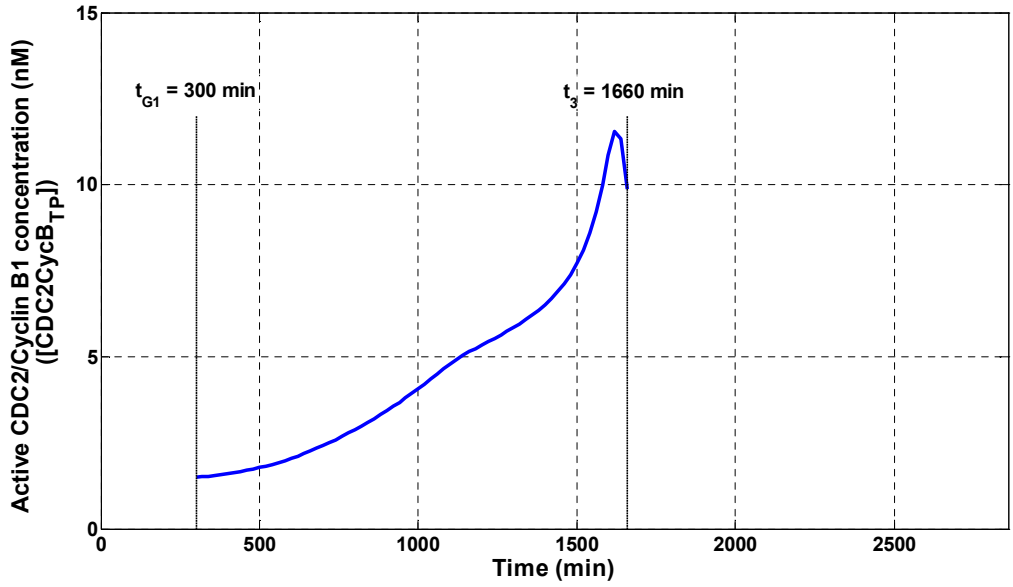


Figure 5.23: Simulation results of the extended cell cycle model Equations (5.1)-(5.11) for the concentration (in nM) of the active form of Cyclin B1/CDC2 complex ($[CDC2CycB_{TP}]$).

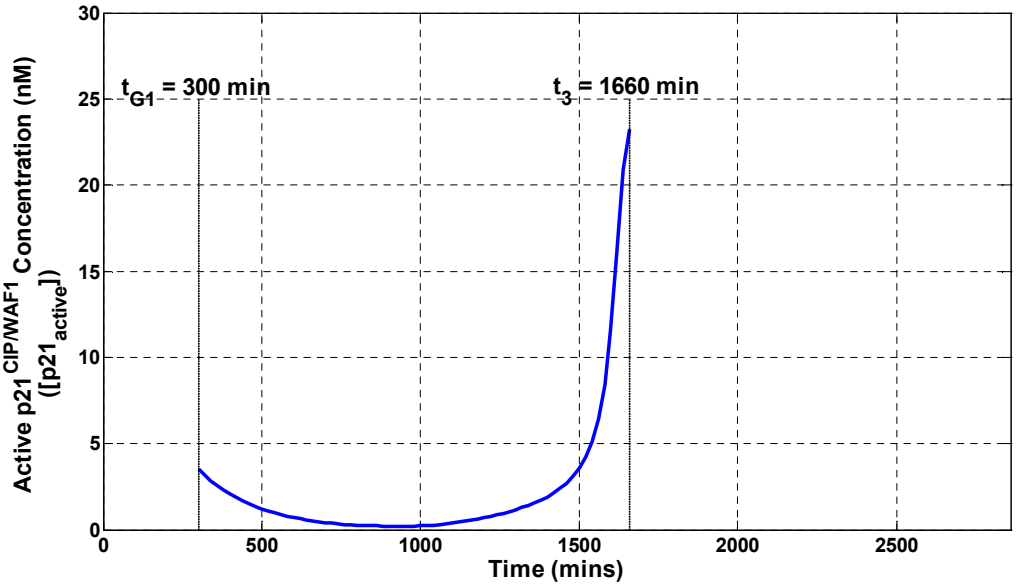


Figure 5.24: Simulation results of the extended cell cycle model Equations (5.1)-(5.11) for the concentration (in nM) of the active form of the CKI $p21^{CIP1/WAF1}$ ($[p21_{active}]$).

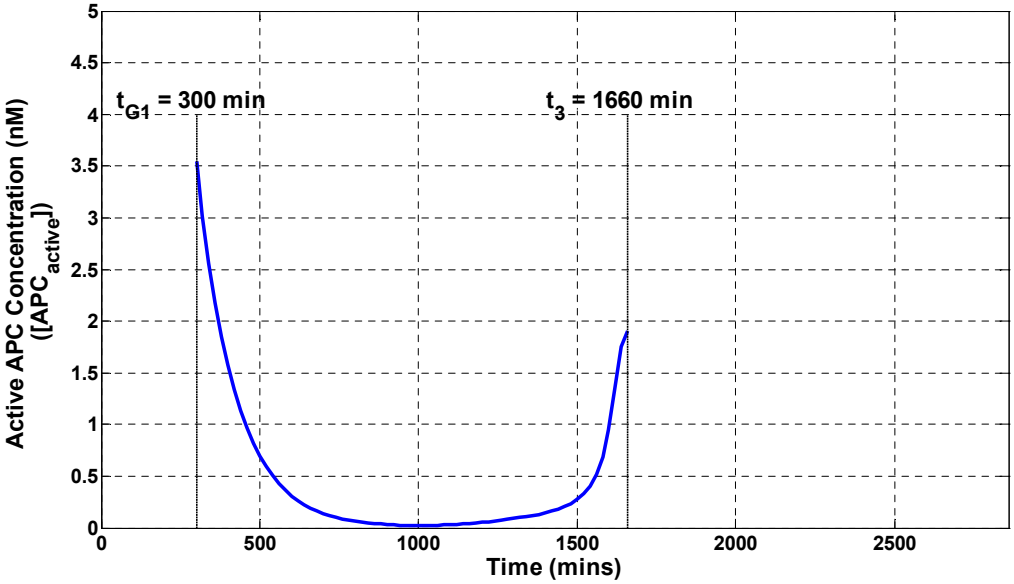


Figure 5.25: Simulation results of the extended cell cycle model Equations (5.1)-(5.11) for the concentration (in nM) of the active form of the APC ($[APC_{active}]$)..

One of the main advantages of mathematical modelling is to infer from the model information relating to variables of the system (in this case, the extended cell cycle model) that are not directly measured (i.e., parts of the system that are not accessible) [13]. For example, it is the fluorescence intensities of the active form of Cyclin B1/CDC2 that are measured using the time-lapse microscopy experiments, and so one might wish to study the changes in the actual concentrations of the cell cycle regulators during the cell cycle.

Figures 5.23, 5.24 and 5.25 represent the concentrations (in nM) of the active forms of Cyclin B1/CDC2 ([$CDC2CycB_{TP}$]) in the cytoplasm, the CKI $p21^{CIP1/WAF1}$ ([$p21_{active}$]) and the APC ([APC_{active}]), respectively, simulated from the extended cell cycle model (Equations (5.1)-(5.11) and Equation (5.16), with parameters taking values in Tables 5.1 and 5.2) during one complete cell cycle (t_{G1} = 300 min to t_3 = 1660 min) of a newborn cell see Figure 5.21. Based on the simulation results presented in Figures 5.23-5.25, at G_1 (at time t_{G1} = 300 min) the concentration of active CDC2 ([$CDC2CycB_{TP}$]) is low as shown in Figure 5.23 while the concentrations of both the APC and the CKI $p21^{CIP1/WAF1}$ are relatively high. During the cell cycle, particularly in the S-G_2-M state (*Start*), the levels of Cyclin B1/CDC2 rapidly rise (since Cyclin B1 synthesis is induced) while the activities of both the APC and the CKI $p21^{CIP1/WAF1}$ are inhibited by active Cyclin B1/CDC2 protein. At approximately time t = 1500 min the G_2/M transition occurs in cell cycle, and the activities of both the APC and the CKI $p21^{CIP1/WAF1}$ start increasing (Figures 5.24 and 5.25). Towards the end of the cell cycle (i.e. at *Finish*) this increase in the activities of APC and $p21^{CIP1/WAF1}$ starts destroying the activity of CDC2 by degrading Cyclin B1 protein resulting in the explosive spikes of CDC2 activity (see Sections 3.3 and 5.3). The results obtained confirm the antagonism between active CDC2 ([$CDC2CycB_{TP}$]) and the APC as well as the antagonism between active CDC2 and the CKI $p21^{CIP1/WAF1}$.

5.6 Sensitivity analysis of the cell cycle model

The impact of varying the values for the rate for turning the CKI $p21^{CIP1/WAF1}$ off (i.e. k_{p21off}) for one complete cell cycle starting at $t = t_{G1}$ of the untreated (control) cell is shown in Figure 5.26. The parameter k_{p21off} was varied in the range of 1×10^{-2} min^{-1}-1×10^{-1} min^{-1}. These values are: (a) 1×10^{-2} min^{-1}, (b) 3×10^{-2} min^{-1}, (c) 5.2973×10^{-2} min^{-1} and (d) 1×10^{-1} min^{-1}. The model variable $CDC2CycB_{TP}$ (the active form of Cyclin B1) exhibits noticeable variation in simulations illustrated in Figure 5.26 when varying the parameter k_{p21off}. That is, the activity of Cyclin B1/CDC2 is sensitive to changes in the rate at which the CKI $p21^{CIP1/WAF1}$ is switched off (deactivated). Based on the results obtained in Figure 5.26, the time required to complete a cell cycle decreases with an increase in the parameter k_{p21off}. This is not surprising since the CKI $p21^{CIP1/WAF1}$ and active CDC2 ($CDC2CycB_{TP}$) are antagonist proteins (see Figures 5.23 and 5.24).

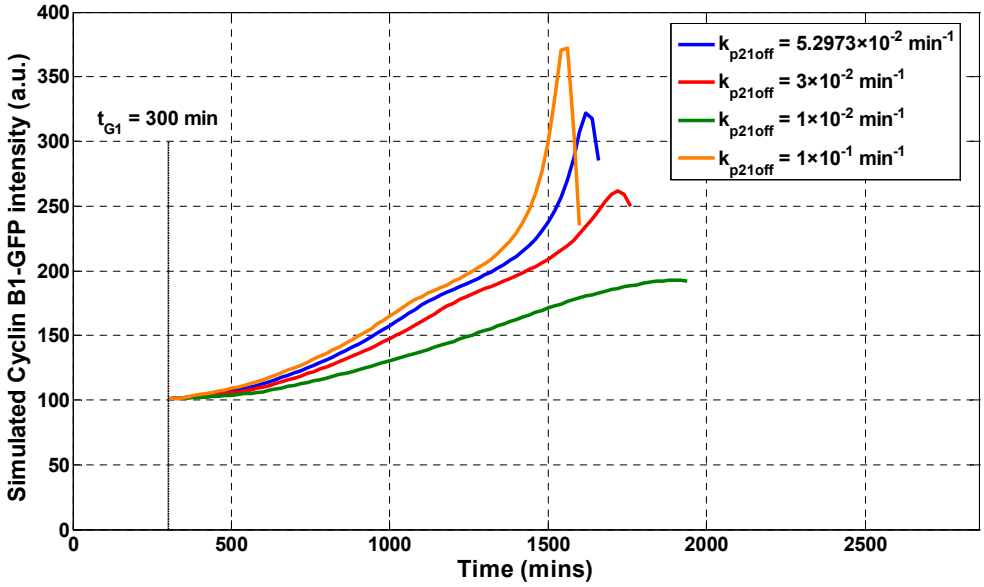


Figure 5.26: Simulated output from the extended cell cycle model Equations (5.1)-(5.11) and state value at $t = t_{G1}$ (Equation (5.16)) by varying rate for turning CKI p21$^{CIP1/WAF1}$ off, $k_{p21off} = 5.2973\times10^{-2}$ min^{-1} (blue), $k_{p21off} = 3\times10^{-2}$ min^{-1} (red), $k_{p21off} = 1\times10^{-2}$ min^{-1} (green) and $k_{p21off} = 1\times10^{-1}$ min^{-1} (orange).

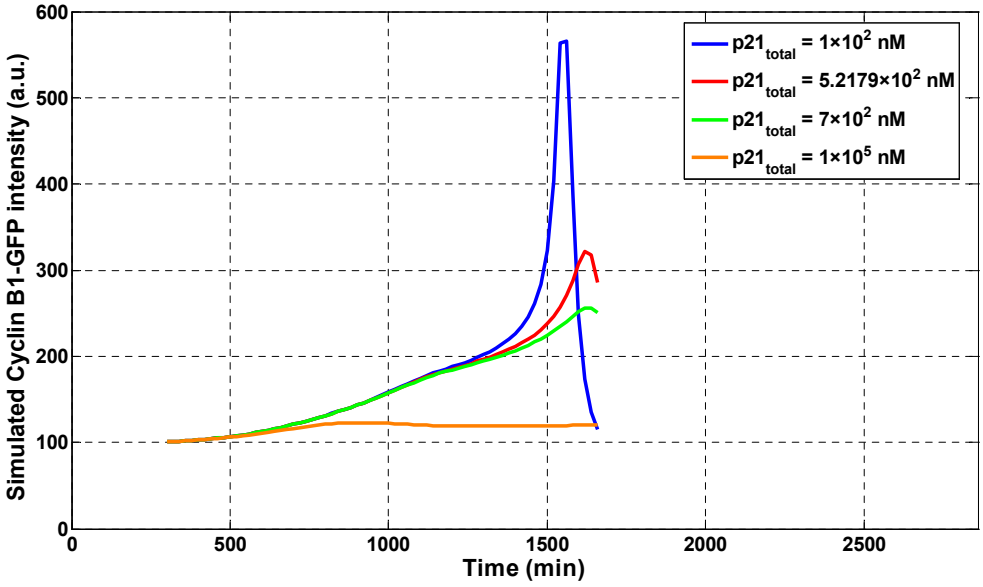


Figure 5.27: Simulated output from the extended cell cycle model Equations (5.1)-(5.11) and state value at $t = t_{G1}$ (Equation (5.16)) by varying the total concentration of CKI , $p21_{total} = 1\times10^{+2}$ nM (blue), $p21_{total} = 5.2179\times10^{+2}$ nM, $p21_{total} = 7\times10^{+2}$ nM and $p21_{total} = 1\times10^{+5}$ nM.

The impact of varying the values for the total concentration of CKI p21$^{CIP1/WAF1}$ (i.e., $p21_{total}$) for one complete cell cycle starting at $t = t_{G1}$ of the untreated (control) cell is shown in Figure 5.27. The parameter $p21_{total}$ was varied in the range of $1\times10^{+2}$ nM-$1\times10^{+5}$ nM. These values are: (a) $1\times10^{+2}$ nM, (b) $5.2179\times10^{+2}$ nM, (c) $7\times10^{+2}$ nM and (d) $1\times10^{+5}$

nM. The model variable $CDC2CycB_{TP}$ (the active form of Cyclin B1) exhibits noticeable variation in simulations shown in Figure 5.27 when varying the parameter $p21_{total}$. That is, the activity of Cyclin B1/CDC2 is sensitive to changes in the total concentration of CKI $p21^{CIP1/WAF1}$. Based on the results obtained in Figure 5.27, the time required to complete a cell cycle increases with an increase in the parameter $p21_{total}$. In addition, the activity of CDC2/Cyclin B1 complex decreases with an increase in the total concentration of CKI $p21^{CIP1/WAF1}$. This antagonist relationship between both CDC2 ($CDC2CycB_{TP}$) and the CKI can result in cell cycle arrest in the G_1 phase of the cell cycle. These results match the expected biological behaviour of cells undergoing a cell cycle.

5.7 Summary

A state-space model describing the response of growth of single human cells in the absence of TPT is presented. The cell cycle model described in this chapter is a *novel* extension of a previously published model by Pomerening *et al.* [14] for *Xenopus laevis* embryos to account for the inhibition effect of the CKI $p21^{CIP1/WAF1}$. The structural identifiability analysis of the extended cell cycle model demonstrates that all of the unknown model parameters are uniquely determined by the output structure corresponding to the real experiment. The structural identifiability analysis of the extended cell cycle model was performed using MATHEMATICA. The unknown model parameters were estimated (using FACSIMILE) from live human osteosarcoma cells (Cyclin B1-GFP data) of a control (untreated cell) for one complete cell cycle event of a newborn cell starting at G_1 of the cycle, to investigate the effects of the CKI $p21^{CIP1/WAF1}$ on the progression of the cell cycle traverse. The model fit of the untreated cell shows that there is close reproduction of the experimental data by simulated output from the extended cell cycle model. The extended cell cycle model is used in the coupled drug kinetics/cell cycle response model.

6 The coupled drug kinetics/dynamics model

6.1 Introduction

This chapter presents a mathematical model describing the response of the growth of single human osteosarcoma cells (U-2 OS cells) to the presence of the anti-cancer agent TPT. The model includes a *novel* coupling of the extended kinetics model of TPT (described in Chapter 4) to the extended cell cycle model (described in Chapter 5). That is, the coupled kinetics/cell cycle response model presented in this thesis describes the effect of the drug on cell cycle dynamics whilst also accounting for the *in vitro* kinetics of the drug (i.e., a description from dose to effect). Coupling the models in this way, rather than studying them separately, allows examination of how the anti-cancer drug TPT perturbs the cell cycle.

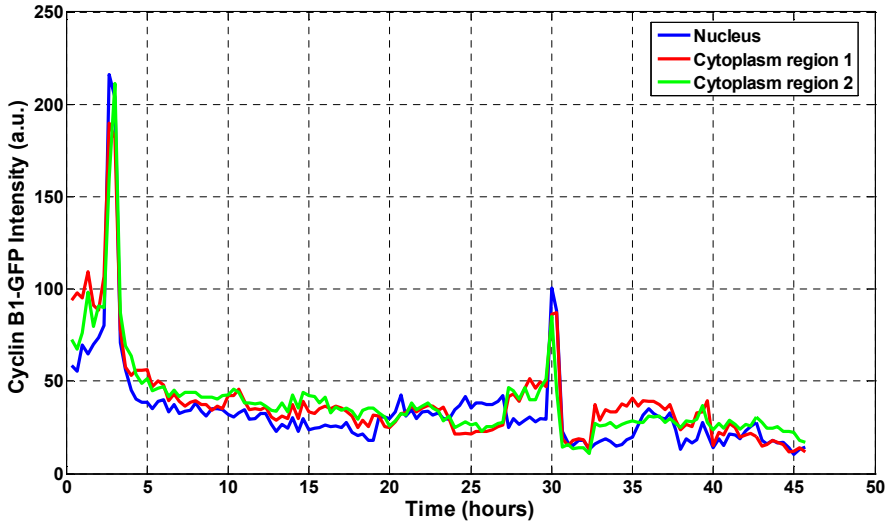
The main feature of the coupled kinetics/cell cycle response (dynamics) model is the ability to demonstrate the dynamic interactions of the active form of the drug (TPT_L) with the DNA-associated molecular target (i.e., the topoisomerase I enzyme, see Section 3.1) and the downstream impact on cell growth (including the cell cycle) and cell death. The coupling between the kinetics model and the cell cycle response (dynamics) model is based on underlying biological assumptions of the cell cycle perturbations induced by the anti-cancer agent TPT. Live-cell data from high-resolution fluorescence imaging of Cyclin B1 (see Sections 5.2 and 6.2) and high-content imaging for quantifying DNA damage (see Section 6.2) have been used for parameter estimation of the unknown model parameters of the proposed coupled kinetics/dynamics model. Moreover, these data are used for model development and parameter estimation in addition to model simulation to further examine the drug interactions with its DNA target and possible routes of cellular evasion of drug action (i.e., drug resistance mechanisms) [107]. Ultimately, a robust and validated kinetics/dynamics model could be used to design and quantitatively predict drug effect (*in silico* predictions) and therefore, the consequences and potential failure of drug treatment regimens (treatment schedules). The analytical and numerical methods described in Chapter 2 are used to investigate the coupled model presented in this chapter.

Generally, the relationships between drug kinetics and the effect of the drug on the cell cycle (i.e., cell cycle response) are exceedingly complicated. In particular, when the perturbed biological system being modelled expresses discrete responses (events) within a heterogeneous population of cells [107]. A model developed by Alarcón *et al.* [271] considers a method that uses the cell cycle (in normal and cancer cells) as a descriptor of the biological response to hypoxia (oxygen deprivation). One limitation of the model described by Alarcón *et al.* [271] is the difficulty in obtaining model validation or comparison with experimental data. More recently, Chappell *et al.* [107] linked the drug kinetic model (for TPT) developed by Evans *et al.* [13] with a basic cell cycle model suggested by Tyson and Novak [157] (for yeast cells). The coupled model (by Chappell *et al.* [107]) is similar to the kinetics/cell cycle response model described in this thesis, in that both models describe the response of the growth of single human osteosarcoma cells (U-2 OS cell line) in the presence of the anti-cancer agent TPT. Moreover, the model developed by Chappell *et al.* [107] has been compared to *in vitro* cell cycle response data similar to the data used in the coupled model described in this chapter. However, a current limitation of the coupled model developed by Chappell *et al.* [107] is that it does not account for drug resistance as a result of BCRP and the ALDH enzyme which affect drug activity and delivery to the nuclear target sites (see Chapter 4). In addition, the cell cycle response model used by Chappell *et al.* [107] does not take into account the effects of the CKI $p21^{CIP1/WAF1}$ which is an important cell cycle regulator in G_1 and G_2 phases of the cell cycle traverse. Moreover, the coupled model described by Chappell *et al.* [107] does not consider the variations in the interaction rate constants between cells (within the same cell type) under different conditions (particularly, different treatment regimens).

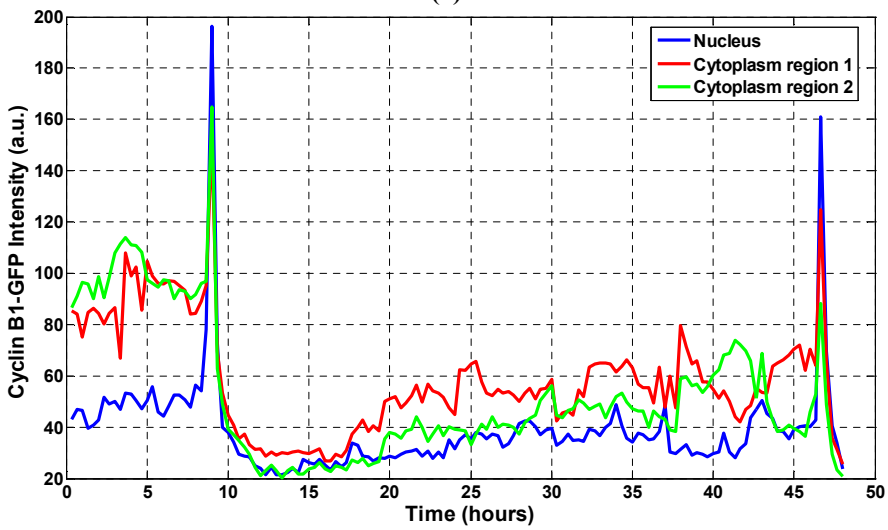
6.2 Experimental data collection

The ability to undertake high spatio-temporal resolution monitoring of cell cycle progression is important and crucial to modelling the sequence of events of cell cycle responses to the action of the anti-cancer agent TPT. This is achieved by the tracking a single cell (U-2 OS cell line) in a non-destructive (non-invasive) fashion within heterogeneous populations using GFP (CCPMs) as explained in Section 5.2. This

monitoring approach also enables continuous tracking of single cell checkpoint transitions non-invasively in the presence of the anti-cancer drug TPT.



(a)



(b)

Figure 6.1: Cyclin B1-GFP intensity profiles extracted from typical G_2 cells treated with (a) 1 µM ($1\times10^{+3}$ nM) TPT and (b) 10 µM ($10\times10^{+3}$ nM) TPT. Three compartments were tracked, the nucleus (blue) and two regions of interest in the cytoplasm (red and green) respectively.

The behaviour of both the progenitor and the descending line of offspring demonstrate the time-integrated biological (cell cycle) response to intervention by the anti-cancer drug TPT including changes in the IMT (illustrated in Figure 6.1) as well as cell death. The time-lapse microscopy experiments provide essential information (Cyclin B1-GFP) upon which the coupled drug kinetics/cell cycle response model is both developed and to a certain

degree validated. These experiments (for tracking cell cycle progression in the presence of TPT) were performed with cells seeded into a multi-well coverslip-bottomed plate. Each well represents a different treatment regimen. That is, live human osteosarcoma cells were treated with 1 μM ($1\times10^{+3}$ nM) or 10 μM ($10\times10^{+3}$ nM) bolus of TPT with 60 min (1 h) residence and then the drug was washed out and the cell tracking began (see Section 6.3).

Similar to the control experiment explained in Section 5.2, Sequences of the fluorescence images were captured every 20 mins in triplicate (i.e., three ROIs) per treatment regime corresponding to one in the nucleus and two in the cytoplasm, as shown in Figure 6.1. Similar to the lineage of the untreated (control) cell (see Chapter 5), a cell lineage for cells treated with TPT illustrates cellular information including phenotypic behaviour as well as fluorescence intensities (Cyclin B1-GFP intensities in a.u.) at the single cell level. Figures 6.1(a) and 6.1(b) show continuous Cyclin B1-GFP tracking extracted for single cells in G_2 phase of the cell cycle traverse treated with 1 μM ($1\times10^{+3}$ nM) and 10 μM ($10\times10^{+3}$ nM) TPT concentrations, respectively. The main effect of treating human osteosarcoma cells with TPT is the extension between the two mitotic landmarks (mitotic events or spikes). that is, an increase in the IMT including an extended delay in the G_2 phase of the cell cycle progression for the cell treated with TPT. Cyclin B1-GFP data (for cells treated with TPT) were used for parameter estimation in Subsection 6.5.2. In addition, the coupled drug kinetics/cell cycle response model in this chapter looks into cells that undergo normal cell division, outcomes including re-fused or polyploidy cells are not in the scope of the work presented in this thesis.

In addition to Cyclin B1-GFP data, the project collaborators at Cardiff University have also provided DNA damage data (using high-content imaging) for live human osteosarcoma cells (U-2 OS cell line) treated with a 1 h bolus of 1 μM ($1\times10^{+3}$ nM) and 10 μM ($10\times10^{+3}$ nM) TPT and then washed out (see Sections 6.3-6.5). Histones are water soluble alkaline proteins around which DNA can wind. These proteins are found in cell nuclei and are highly involved in gene regulation in eukaryotic cells by packaging and ordering DNA into the basic structural units called nucleosomes [272]. Histone H2AX is one of several variants of histone H2A (one of the six histone classes) [273-275], that is immediately phosphorylated at its C terminus on serine 139 by the Ataxia telangiectasia mutated (ATM) protein kinase (when activated) in response to damage that induces dsDNA breaks [276,

277]. Moreover, the phosphorylated H2AX, also termed gamma-H2AX (γH2AX), plays a crucial role in safeguarding genome integrity [278].

ATM protein kinase is also required for the maintenance of genome integrity and is the vital component of the signal transduction pathways that are mobilised by the damage of DNA structure [278]. In addition to H2AX, the p53 tumour suppressor gene (a nucleoplasmic substrate) is among the downstream target effectors phosphorylated by ATM protein kinase. This phosphorylation leads to up-regulation (transcription) of $p21^{CIP1/WAF1}$ protein, thereby, causing cell cycle arrest in the G_1 stage of the cell cycle traverse (see Section 3.3). The phosphorylation of the ATM downstream substrates (including H2AX and p53) is essential for successful DNA repair and suppression of tumourigenesis (formation of tumours). That is, suppressing cells from further progression through the cell cycle and/or inducing apoptosis (cell death) and therefore, preventing transmission of DNA damage to cell progeny that could lead to genome instability (i.e., to enable cells to accumulate mutations to develop malignancies).

ATM activation and H2AX phosphorylation, consequences of DNA damage, can either be induced by radiation, anti-cancer drugs or environmental carcinogens/mutagens. However, the scope of this chapter is to asses the response of the growth (including DNA damage) of single human cells induced by the DNA topoisomerase I inhibitor TPT. This anti-cancer agent (i.e., TPT) is involved in stabilising the cleavable complex formed between topoisomerase I enzyme which collides with a progressing DNA replication fork generating irreversible dsDNA breaks (see Section 3.1). These breaks in chromatin, (i.e., dsDNA breaks) induce a cascade of molecular events to sense, transduce and effect DNA repair and cellular responses.

Expression of ATM-mediated γH2AX foci act as a reporter of DNA damage that generates formation of dsDNA breaks [279, 280]. In addition, the number of induced dsDNA breaks is reflected by the frequencies of γH2AX foci [281]. Therefore, the detection of γH2AX foci by imaging provides a quantitative tool for determining the action of the anti-cancer agent TPT that induces DNA damage and potentially enables the linking between DNA damage and repair with the initiation of the cell cycle checkpoints, delay and arrest. The

high-content imaging provides the final number (count) of γH2AX foci in single nuclei within a heterogeneous population of live human osteosarcoma cells (U-2 OS cell line) treated with 1 μM ($1\times10^{+3}$ nM) or 10 μM ($10\times10^{+3}$ nM) bolus of TPT for 1 h.

6.3 The coupled drug kinetics/cell cycle model

Figure 6.2 provides a schematic for the coupled drug kinetic/cell cycle model. Separate models have been developed for the drug kinetics (see Chapter 4) and the cell cycle (see Chapter 5). These models are: the extended state-space kinetic model that describes the drug activity and delivery of the pharmacologically active form (TPT_L) to the DNA target as well the catalysis of the ALDH enzyme and the elimination of drug from the cytoplasm via the efflux pump (BCRP/ABCG2 transporter); and the extended state-space model that describes the continuous tracking of Cyclin B1 through the cell cycle at the single level, to account for the inhibition of $p21^{CIP1/WAF1}$ protein on the dynamics of Cyclin B1.

In this coupled model, relevant basic physiology and biochemical mechanisms are translated into a system of non-linear ODEs by the law of mass action (see Equations (6.1)-(6.15)). That is, the mathematical modelling of the coupled drug kinetics/cell cycle model involves the construction of ODEs to describe protein interaction networks that regulate the cell cycle as well as the chemical reactions, simple diffusion, active pumping mechanism (BCRP/ABCG2 transporter) and partitions that take place when the drug moves to/from the medium, the cytoplasm and nuclear compartments. The coupling of both models (i.e., the cell cycle model and the drug kinetics model) allows the analysis of how the drug perturbs and affects cell cycle progression in live human osteosarcoma cells (the U-2 OS cell line).

The method of modelling the dynamic effect (response) of TPT on human osteosarcoma cells is based on the approach described in Subsection 2.3.4. The pseudo-steady state approximated drug kinetics model in Figure 4.9 has been extended by introducing an effect (response to TPT) compartment denoted by the variable $\gamma H2AX(t)$ with rate constants k_{TPT} and k_{dam} (see Subsection 2.3.4). The variable $L_n(t)$ remains unchanged by the addition of

the extra compartment. The new variable, (i.e., $\gamma H2AX(t)$), represents the number of γH2AX foci (i.e., foci count) in a single nucleus. The set of ODEs describing the drug kinetics part (including the dynamic effect) of the coupled model is given by:

$$\begin{aligned}
\frac{\mathrm{d}L_m}{\mathrm{dt}} &= -(k_{om}+k_{mi})L_m + k_{cm}H_m + k_{mo}v_0L_e \\
\frac{\mathrm{d}H_m}{\mathrm{dt}} &= k_{om}L_m - (k_{cm}+k_{mi})H_m + k_{mo}v_0H_e \\
\frac{\mathrm{d}L_e}{\mathrm{dt}} &= \frac{k_{mi}L_m}{v_0} - (k_{mo}+k_{om}+k_i)L_e + k_{cm}H_e + \frac{k_eL_c}{v_1} \\
\frac{\mathrm{d}H_e}{\mathrm{dt}} &= \frac{k_{mi}}{v_0}H_m + k_{om}L_e - (k_{cm}+k_{mo})H_e + \frac{v_3V_{m1}H_c}{v_1(v_3k_{m1}+H_c)} \\
\frac{\mathrm{d}L_c}{\mathrm{dt}} &= k_iv_1L_e - (k_e+k_{oc})L_c + k_{cc}H_c + k_{dl}v_2L_n - k_b(B_T-L_n)L_c \\
&\quad - \frac{V_{\mathrm{m}2}L_c}{(k_{m2}+L_c)} \\
\frac{\mathrm{d}H_c}{\mathrm{dt}} &= k_{oc}L_c - k_{cc}H_c + k_{dh}v_2L_n - \frac{v_3V_{m1}H_c}{(v_3k_{m1}+H_c)} + \frac{V_{m2}L_c}{(k_{m2}+L_c)} \\
\frac{\mathrm{d}L_n}{\mathrm{dt}} &= \frac{k_b}{v_2}(B_T-L_n)L_c - (k_{dl}+k_{dh})L_n \\
\frac{\mathrm{d}\gamma H2AX}{\mathrm{dt}} &= k_{\mathrm{TPT}}L_n - k_{dam}\gamma H2AX
\end{aligned} \tag{6.1}$$

where, $v_0 = V_e/V_m$, $v_1 = V_e/V_c$, $v_2 = V_n/V_c$, $v_3 = V_{cm}/V_c$, and $v_1/v_3 = V_e/V_{cm}$. The volumes of the medium, extracellular region, cytoplasm, cell membrane and nucleus, are denoted by V_m, V_e, V_c, V_{cm}, and V_n, respectively (see Section 4.3).

In the drug kinetics model described in Chapter 4, the unknown model parameters were estimated from live human breast cancer cells (MCF-7 cell line) and the coupled model describes the response of growth of the growth of single human osteosarcoma cells (U-2 OS cells) in the presence of TPT. Therefore, for simplicity, it is assumed that cellular volumes ($V_n + V_c + V_{cm}$) are the same across cell types. In addition, the total volume of the medium, V_T, is equal to the sum of volumes of the medium region and the extracellular region, that is,

$$V_T = V_m + V_e = \frac{V_e}{v_0} + V_e = \frac{V_e(1+v_0)}{v_0}. \tag{6.2}$$

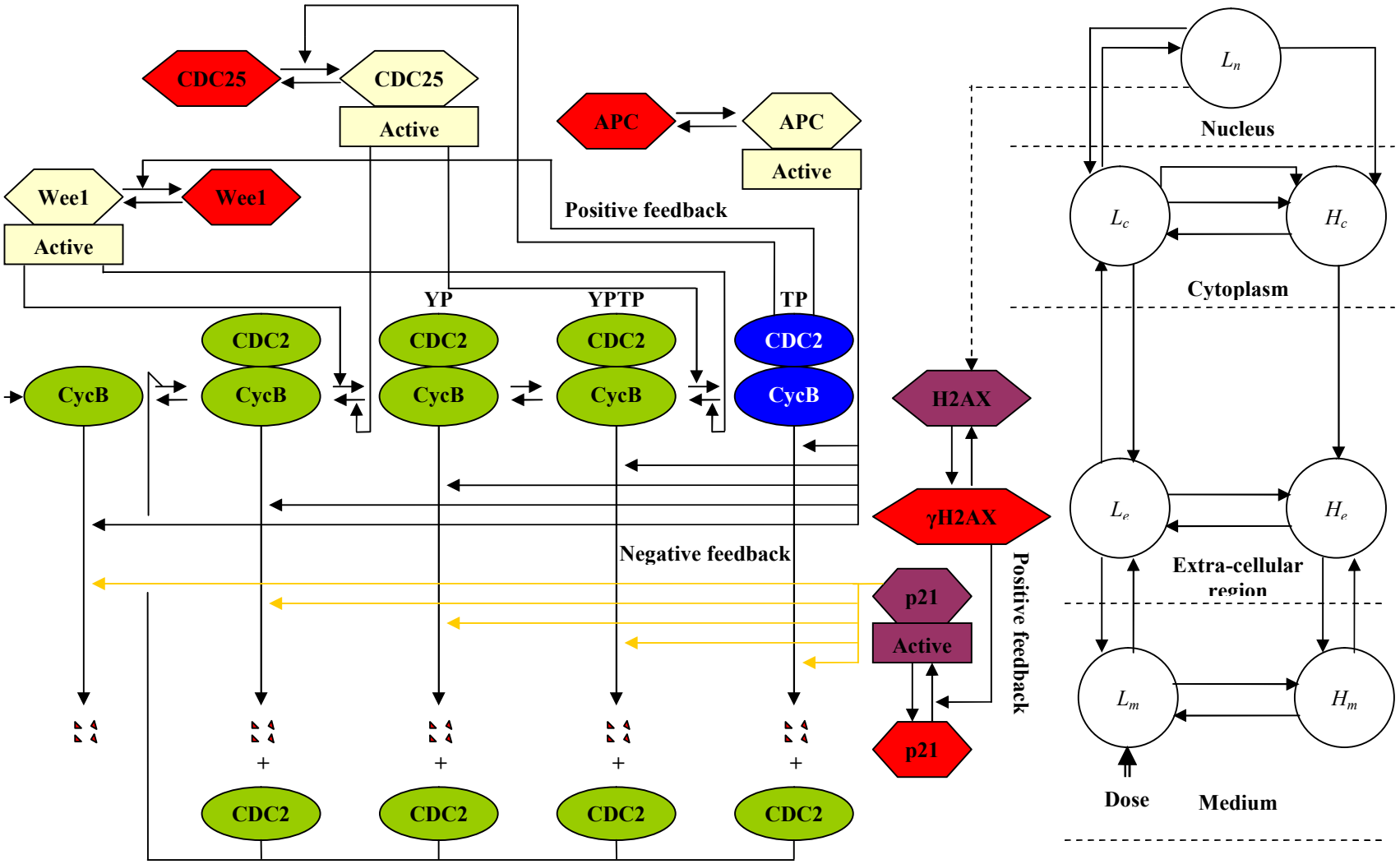


Figure 6.2: Schematic diagram of the coupled drug kinetics/cell cycle response model.

In the coupled drug kinetics/cell cycle response model, it is assumed that DNA damage (represented by γH2AX) is only induced in the presence of TPT, i.e., there is no DNA damage prior to drug administration. Time $t = 0$ corresponds to the (first) addition of TPT_L only as a bolus injection to give a concentration of D (nM) with respect to the total volume of the medium ($V_T = 2$ ml or 2×10^{12} μm^3). The corresponding initial conditions for the drug kinetics model are given by:

$$\begin{aligned} &H_m(0) = L_e(0) = H_e(0) = L_c(0) = H_c(0) = L_n(0) = \gamma H2AX(0) = 0, \text{ and} \\ &L_m(0) = (1+v_0)D. \end{aligned} \tag{6.3}$$

Cells were placed in a medium with pH = 7.2 in the experiments used to collect data for estimating the unknown parameters in the proposed drug kinetics model (see Chapter 4). The parameter values used in the kinetics component (in this chapter) of the coupled drug kinetics/cell cycle model are the estimated parameter values for averaged data from the 13 individual cells (see Table 4.4). After 1 h, the medium containing human osteosarcoma cells is exchanged for drug-free medium. This process is termed *washout* and the cell tracking (Cyclin B1-GFP experiment) begins (20 min after the *washout* event). To model the *washout* at $t = 1$ h the variables relating to the drug (both forms TPT_L and TPT_H) in the medium (medium environment and extracellular region) are instantaneously set to 0, that is:

$$L_m(1) = H_m(1) = L_e(1) = H_e(1) = 0 \tag{6.4}$$

All other variables, i.e. $L_c(t)$, $H_c(t)$, $L_n(t)$ and $\gamma H2AX(t)$ within the model described by Equations (6.1) are unaffected by the *washout* event. The values of the variables relating to the drug (TPT_L and TPT_H) in the cytoplasm and the nucleus, namely, $L_c(t)$, $H_c(t)$, and $L_n(t)$, as well as the value of the variable relating to DNA damage, $\gamma H2AX(t)$, at $t = 1$ h depend on the treatment regimen (i.e., concentration of the dose D), the method for obtaining these values is explained in Section 6.5.

The cellular response to DNA damage involves a series of events including cell cycle arrest (in G_1 phase) or apoptosis. γH2AX foci play an important role in the DNA damage signalling pathway induced by dsDNA breaks by acting as docking areas for the recruitment of DNA repair factors by bringing the broken ends of the DNA closer to

complete the repair. Moreover, Fernandez-Capetillo *et al.* [282] and Fragkos *et al.* [283] have demonstrated the direct significance of γH2AX in cell cycle arrest. That is, γH2AX is necessary for $p21^{CIP1/WAF1}$-induced cell cycle arrest at G_1 phase (i.e. up-regulation of $p21^{CIP1/WAF1}$) after replication stalling [283]. The loss of the CKI $p21^{CIP1/WAF1}$ results in replication defects and this may lead to cell death when cells are treated with anti-cancer agents, including TPT [284, 285]. Proteosome-dependent degradation of $p21^{CIP1/WAF1}$ follows $p53/p21^{CIP1/WAF1}$-mediated cell cycle arrest (in G_1 phase), this process (i.e., degradation of the CKI) results in cells progressing to the S phase of the cell cycle process.

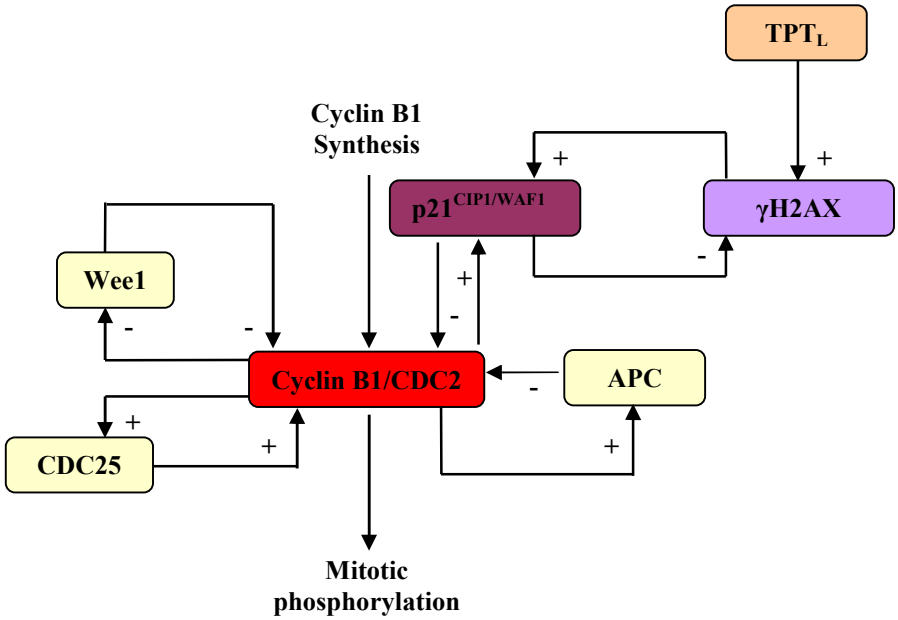


Figure 6.3: Schematic depiction of the coupled drug kinetics/cell cycle response model.

The formation of γH2AX (as a result of DNA damage) has been identified across species from humans to yeast including *Xenopus laevis* embryos [283, 286]. Therefore, the drug kinetics model (presented in Chapter 4) is coupled to the extended cell cycle model (presented in Chapter 5) as shown in Figures 6.2 and 6.3. The drug kinetics model has been linked to the CKI and regulator $p21^{CIP1/WAF1}$ via γH2AX, which is itself driven by the kinetic model within the cell cycle constraint. This coupling between the two models involves an additional circuit element, a negative-feedback component to account for the regulation between the DNA damage induced by TPT which is represented by γH2AX and the CKI $p21^{CIP1/WAF1}$. These feedback loops could function as a bistable system (switch) with a hysteretic steady-state stimulus/response relationship, toggling between two discrete

steady states, namely, G_1 state and S-G_2-M state (see Chapter 5). In the model presented in this chapter, the coupling between γH2AX and bound TPT in the nucleus is constrained to only have an effect in the S-G_2-M state of the cell cycle process since TPT is believed to be specific to the S phase (see Section 3.3) and therefore, TPT induced γH2AX acts almost exclusively in S phase cells [278].

Figure 6.3 provides a schematic diagram of the structure of the coupled drug kinetic/cell cycle response model. The model equations are based upon those of the cell cycle model describing the mechanisms (biochemical reactions) occurring in untreated cells (see Chapter 5). The system of equations employed in the coupled drug kinetics/cell cycle response model is a modified version of these based upon the action of TPT and using appropriate parameter values derived from live-cell data as explained in Section 6.2 (the human osteosarcoma cells). The equations (processes) governing the biochemical interactions during the cell cycle process, in the presence of TPT, are given by:

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CycB}]}{\mathrm{dt}} &= C_1 k_{synth} - C_1 k_{dest}([\mathrm{APC}_{\mathrm{active}}]+[\mathrm{p21}_{\mathrm{active}}])[\mathrm{CycB}] \\ &\quad - C_1 k_a[\mathrm{CDC2}][\mathrm{CycB}] + C_1 k_d[\mathrm{CDC2CycB}]\end{aligned} \tag{6.5}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2}]}{\mathrm{dt}} &= C_1 k_{dest}([\mathrm{APC}_{\mathrm{active}}]+[\mathrm{p21}_{\mathrm{active}}])([\mathrm{CDC2CycB}] \\ &\quad + [\mathrm{CDC2CycB}_{\mathrm{YPTP}}]+[\mathrm{CDC2CycB}_{\mathrm{YP}}] \\ &\quad + [\mathrm{CDC2CycB}_{\mathrm{TP}}]) \\ &\quad - C_1 k_a[\mathrm{CDC2}][\mathrm{CycB}] + C_1 k_d[\mathrm{CDC2CycB}]\end{aligned} \tag{6.6}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2CycB}]}{\mathrm{dt}} &= C_1 k_a[\mathrm{CDC2}][\mathrm{CycB}] - C_1 k_d[\mathrm{CDC2CycB}] \\ &\quad - C_1 k_{dest}([\mathrm{APC}_{\mathrm{active}}]+[\mathrm{p21}_{\mathrm{active}}])[\mathrm{CDC2CycB}] \\ &\quad - C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1}_{\mathrm{active}}][\mathrm{CDC2CycB}] \\ &\quad \frac{-C_1 C_2 k_{\mathrm{Wee1}}[\mathrm{Wee1}_{\mathrm{total}}][\mathrm{CDC2CycB}]}{factor} \\ &\quad \frac{+C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1}_{\mathrm{active}}][\mathrm{CDC2CycB}]}{factor} \\ &\quad + C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{YP}}] \\ &\quad \frac{+C_1 C_2 k_{\mathrm{CDC25}}[\mathrm{CDC25}_{\mathrm{total}}][\mathrm{CDC2CycB}_{\mathrm{YP}}]}{factor} \\ &\quad \frac{-C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{YP}}]}{factor}\end{aligned} \tag{6.7}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2CycB_{YP}}]}{\mathrm{dt}} &= C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB}] \\ &- C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YP}}] \\ &- C_1 k_{\mathrm{CAK}}[\mathrm{CDC2CycB_{YP}}] + C_1 k_{\mathrm{PP2C}}[\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{dest}([\mathrm{APC_{active}}]+[\mathrm{p21_{active}}])[\mathrm{CDC2CycB_{YP}}] \\ &\frac{+C_1 C_2 k_{\mathrm{Wee1}}[\mathrm{Wee1_{total}}][\mathrm{CDC2CycB}]}{factor} \\ &\frac{-C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB}]}{factor} \\ &\frac{-C_1 C_2 k_{\mathrm{CDC25}}[\mathrm{CDC25_{total}}][\mathrm{CDC2CycB_{YP}}]}{factor} \\ &\frac{+C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YP}}]}{factor}\end{aligned} \tag{6.8}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2CycB_{YPTP}}]}{\mathrm{dt}} &= C_1 k_{\mathrm{CAK}}[\mathrm{CDC2CycB_{YP}}] - C_1 k_{\mathrm{PP2C}}[\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}] \\ &+ C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}] \\ &- C_1 k_{dest}([\mathrm{APC_{active}}]+[\mathrm{p21_{active}}])[\mathrm{CDC2CycB_{YPTP}}] \\ &\frac{-C_1 C_2 k_{\mathrm{CDC25}}[\mathrm{CDC25_{total}}][\mathrm{CDC2CycB_{YPTP}}]}{factor} \\ &\frac{+C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}]}{factor} \\ &\frac{+C_1 C_2 k_{\mathrm{Wee1}}[\mathrm{Wee1_{total}}][\mathrm{CDC2CycB_{TP}}]}{factor} \\ &\frac{-C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}]}{factor}\end{aligned} \tag{6.9}$$

$$\begin{aligned}\frac{\mathrm{d}[\mathrm{CDC2CycB_{TP}}]}{\mathrm{dt}} &= C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}] \\ &- C_1 k_{dest}([\mathrm{APC_{active}}]+[\mathrm{p21_{active}}])[\mathrm{CDC2CycB_{YPTP}}] \\ &\frac{+C_1 C_2 k_{\mathrm{CDC25}}[\mathrm{CDC25_{total}}][\mathrm{CDC2CycB_{YPTP}}]}{factor} \\ &\frac{-C_1 k_{\mathrm{CDC25}}[\mathrm{CDC25_{active}}][\mathrm{CDC2CycB_{YPTP}}]}{factor} \\ &- C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}] \\ &\frac{-C_1 C_2 k_{\mathrm{Wee1}}[\mathrm{Wee1_{total}}][\mathrm{CDC2CycB_{TP}}]}{factor} \\ &\frac{+C_1 k_{\mathrm{Wee1}}[\mathrm{Wee1_{active}}][\mathrm{CDC2CycB_{TP}}]}{factor}\end{aligned} \tag{6.10}$$

$$\frac{\mathrm{d}[\mathrm{APC}_{\mathrm{active}}]}{\mathrm{dt}} = \frac{C_1 C_2 k_{\mathrm{APCon}}[\mathrm{APC}_{\mathrm{total}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{APC}}}}{(C_2[\mathrm{EC50}_{\mathrm{APC}}])^{n_{\mathrm{APC}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{APC}}}} - \frac{C_1 k_{\mathrm{APCon}}[\mathrm{APC}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{APC}}}}{(C_2[\mathrm{EC50}_{\mathrm{APC}}])^{n_{\mathrm{APC}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{APC}}}} - C_1 k_{\mathrm{APCoff}}[\mathrm{APC}_{\mathrm{active}}] \tag{6.11}$$

$$\frac{\mathrm{d}[\mathrm{Wee1}_{\mathrm{active}}]}{\mathrm{dt}} = \frac{-C_1 k_{\mathrm{Wee1off}}[\mathrm{Wee1}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{Wee1}}}}{(C_2[\mathrm{EC50}_{\mathrm{Wee1}}])^{n_{\mathrm{Wee1}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{Wee1}}}} + C_1 k_{\mathrm{Wee1on}}(C_2[\mathrm{Wee1}_{\mathrm{total}}] - [\mathrm{Wee1}_{\mathrm{active}}]) \tag{6.12}$$

$$\frac{\mathrm{d}[\mathrm{CDC25}_{\mathrm{active}}]}{\mathrm{dt}} = \frac{C_1 C_2 k_{\mathrm{CDC25on}}[\mathrm{CDC25}_{\mathrm{total}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}}{(C_2[\mathrm{EC50}_{\mathrm{CDC25}}])^{n_{\mathrm{CDC25}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}} - \frac{C_1 k_{\mathrm{CDC25on}}[\mathrm{CDC25}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}}{(C_2[\mathrm{EC50}_{\mathrm{CDC25}}])^{n_{\mathrm{CDC25}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{CDC25}}}} - C_1 k_{\mathrm{CDC25off}}[\mathrm{CDC25}_{\mathrm{active}}] \tag{6.13}$$

$$\frac{\mathrm{d}[\mathrm{p21}_{\mathrm{active}}]}{\mathrm{dt}} = \frac{C_1 C_2 k_{\mathrm{p21on}}[\mathrm{p21}_{\mathrm{total}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}}{(C_2[\mathrm{EC50}_{\mathrm{p21}}])^{n_{\mathrm{p21}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}} - \frac{C_1 k_{\mathrm{p21on}}[\mathrm{p21}_{\mathrm{active}}][\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}}{(C_2[\mathrm{EC50}_{\mathrm{p21}}])^{n_{\mathrm{p21}}} + [\mathrm{CDC2CycB}_{\mathrm{TP}}]^{n_{\mathrm{p21}}}} - \frac{C_1 k_{\mathrm{p21off}}[\mathrm{p21}_{\mathrm{active}}]}{(1 + k_{dam}\gamma H2AX)} . \tag{6.14}$$

In the coupled drug kinetics/cell cycle response model, it is assumed that DNA damage (represented by γH2AX) is only induced in the presence of TPT. The up-regulation of the CKI $p21^{CIP1/WAF1}$ resulting from the increase in γH2AX foci has been incorporated in the model by decreasing the rate constant at which $p21^{CIP1/WAF1}$ is turned off (i.e., k_{p21off}). This has been employed in the model by multiplying k_{p21off} by the factor of $1/(1 + k_{dam}\gamma H2AX)$ as shown in Equation (6.14). Accordingly, this increases the time duration of the CKI's activity, in addition, in the absence of TPT, the equation that relates to the activity of $p21^{CIP1/WAF1}$ reduces to Equation (5.10) (see Chapter 5) since the number of γH2AX foci reduces to zero.

Similar to the cell cycle model described in Chapter 5, the scaling factors C_1 and C_2 (unknown model parameters) are the scaling factors introduced to account for the

variations in the rate constants and variations in the concentrations of the cell cycle components among different organisms, respectively. Table 6.1 represents the values for the parameters of the coupled drug kinetics/cell cycle response model which are characterised by those given in Tables 4.4 (with appropriate conversion of units) and 5.1. In this chapter, time $t = 0$ min corresponds to the (first) addition of TPT_L (only) as a bolus injection. The medium is exchanged for drug-free medium (i.e., *washout* event) at time $t_w = 60$ min. In addition, the corresponding values for the cell cycle response model variables at the start of a new cell cycle at time $t = t_{G1}$ are given by (see Section 5.3):

$$\begin{gathered}
[\text{CycB}](t_{G1}) = 0.141396C_2, [\text{CDC2CycB}](t_{G1}) = 0.992654C_2, \\
[\text{CDC2CycB}_{\text{YP}}](t_{G1}) = 1.4763C_2, [\text{CDC2CycB}_{\text{YPTP}}](t_{G1}) = 11.41023C_2, \\
[\text{CDC2CycB}_{\text{TP}}](t_{G1}) = 1.64121C_2, [\text{CDC2}](t_{G1}) = 84.479606C_2, \\
[\text{APC}_{\text{active}}](t_{G1}) = 3.8539C_2, [\text{Wee1}_{\text{active}}](t_{G1}) = 15C_2, \\
[\text{CDC25}_{\text{active}}](t_{G1}) = 1.54009C_2 \text{ and } [\text{p21}_{\text{active}}](t_{G1}) = 3.8539C_2
\end{gathered} \quad . \qquad (6.15)$$

In all of the time-lapse microscopy experiments (from the U-2 OS cell line) considered in this chapter, fluorescence intensities (which reflect the concentrations of the active form of Cyclin B1, i.e., $[\text{CDC2CycB}_{\text{TP}}]$) are measured in arbitrary units (a.u.) and time in minutes (min). Moreover, in the high-content screening imaging experiments (used to detect γH2AX), DNA damage is represented with the number (count) of γH2AX foci (from U-2 OS cells) and time in minutes. For completeness, the fifty four parameters considered in the coupled drug kinetics/dynamics model are:

- k_{om} and k_{cm}, units: min^{-1} (per minute). Rate constants for deactivation/activation of TPT_L/ TPT_H, respectively, in the medium and extracellular region;
- k_{mi} and k_{mo}, units: min^{-1}. Rate constants for the flow of TPT_L and TPT_H to/from the extracellular region/medium respectively;
- k_i and k_e, units: min^{-1}. Rate constants for the flow of TPT_L into and out of the cell (i.e., to/from extracellular region), respectively;
- k_{oc} and k_{cc}, units: min^{-1}. Rate constants for deactivation/activation of TPT_L/ TPT_H, respectively, in the cell (cytoplasm);
- k_b, units: $(\text{nM min})^{-1}$. Rate constant for binding (association) of TPT_L (in the cytoplasm) to DNA;

- k_{dl} and k_{dh}, units: min^{-1}. Rate constants of dissociation of TPT_L from DNA to TPT_L/ TPT_H (in cytoplasm) respectively;
- B_T units: nM. Concentration of DNA sites available for TPT_L (in cytoplasm) to bind to;
- k_{m1} and k_{m2}, units: nM. Michaelis-Menten constants for the BCRP/ABCG2 drug transporter and the ALDH enzyme respectively.
- V_{m1} and V_{m2}, units: nM min^{-1}. Numerical constants that represent the maximum velocity obtained when the transporter and the enzyme exist completely in the form TH_c and EL_c respectively;
- v_0, v_1, v_2 and v_3. The ratios, respectively, of the volumes of extracellular region (V_e) to medium (V_m), extracellular region to cytoplasm (V_c), nucleus (V_n) to cytoplasm and cell membrane (V_{cm}) to cytoplasm;
- k_{synth} and k_{dest}, units: nM min^{-1} (nM per minute) and $(nM\ min)^{-1}$, respectively. Rate constants for Cyclin B1 (monomer) synthesis and Cyclin B1 (monomer and dimers) destruction, respectively;
- k_a and k_d, units: $(nM\ min)^{-1}$ and min^{-1}, respectively. Rate constants for association of CDC2 (monomer) to Cyclin B1 (monomer) forming the inactive non-phosphorylated complex (CDC2CycB) and the dissociation of this complex to the monomers (i.e., Cyclin B1 and CDC2), respectively;
- *factor*, dimensionless constant. This parameter represents the strength of the feedback from CDC2 to Wee1 kinase and CDC25 phosphatase;
- k_{Wee1}, units: $(nM\ min)^{-1}$. Rate constant for the inactivation of Cyclin B1/CDC2 dimers, namely, the non-phosphorylated complex (CDC2CycB) and the active Cyclin B1/CDC2 complex ($CDC2CycB_{TP}$) by Wee1 kinase;
- k_{CDC25}, units: $(nM\ min)^{-1}$. Rate constant for the activation of Cyclin B1/CDC2 dimers, namely, the Y15-phosphorylated complex ($CDC2CycB_{YP}$) and the Y15- and T161-phosphorylated complex ($CDC2CycB_{YPTP}$) by CDC25 phosphatase;
- $CDC2_{total}$, APC_{total}, $Wee1_{total}$, $CDC25_{total}$ and $p21_{total}$, units: nM. The total concentrations of CDC2, APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;

- n_{APC}, n_{Wee1}, n_{CDC25}, and n_{p21}, dimensionless constants. These dimensionless constants are the Hill coefficients of the Hill functions describing the rates of change of the active forms of the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- $EC50_{APC}$, $EC50_{Wee1}$, $EC50_{CDC25}$, and $EC50_{p21}$, units: nM. The half maximal effective concentrations (EC_{50}'s) in the Hill functions describing the rates of change of the active forms of the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- k_{APCon}, k_{Wee1on}, $k_{CDC25on}$ and k_{p21on}, units: min^{-1}. Rate constants for turning on the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- k_{APCoff}, $k_{Wee1off}$, $k_{CDC25off}$ and k_{p21off}, units: min^{-1}. Rate constants for turning off the APC, Wee1 kinase, CDC25 phosphatase and the CKI $p21^{CIP1/WAF1}$, respectively;
- k_{CAK} and k_{PP2C}, units: min^{-1}. The rate constants for phosphorylation/ dephosphorylation of the Y15-phosphorylated complex ($CDC2CycB_{YP}$)/ Y15- and T161-phosphorylated complex ($CDC2CycB_{YPTP}$), respectively.
- C_1 and C_2, dimensionless constants (parameters). Scaling factors to account for the variations in the interaction rate constants and the variations in the concentrations of the cell cycle components (among different organisms), respectively.
- k_{TPT} and k_{dam}, units: $(nM\ min)^{-1}$ and min^{-1}, respectively. Rate constants relating to the effect (response to TPT) compartment.

Five of these parameters (k_{TPT}, k_{dam}, C_2, k_{p21off} and $p21_{total}$) in addition to two observational parameters (α and β, see Sections 5.4 and 6.4.2) are assumed to be unknown, and need to be estimated from experimental data. Using high-content imaging (for cells treated with $1\times10^{+3}$ and $10\times10^{+3}$ nM TPT), it is possible to detect and count γH2AX foci in single nuclei, and data from these experiments were used to estimate the unknown model parameters k_{TPT} and k_{dam} prior to the *washout* event (at $t \leq 1$ h). The time-lapse microscopy experiments allow direct measurement of the fluorescence intensities (Cyclin B1-GFP) of concentrations of the active form of Cyclin B1 (Cyclin B1/CDK1) in the cytoplasm and nucleus. These experiments were used to estimate the unknown model parameters k_{TPT}, k_{dam}, C_2, k_{p21off} and $p21_{total}$. Prior to the parameter estimation for the coupled drug

kinetics/cell cycle model, two structural identifiability analyses have been performed (see Section 6.4) to test whether the unknown model parameters are uniquely determined by the observations corresponding to the two different experiments. In the following section, the Taylor series approach is used to show that the coupled drug kinetics/cell cycle response model is structurally globally identifiable.

Table 6.1: Table of known parameter values in the drug kinetics model (see Chapter 4) and cell cycle model (see Chapter 5).

Drug kinetics model		Cell cycle model	
Parameter	**Value**	**Parameter**	**Value**
k_i (min^{-1})	1.4159	k_{synth} (nM min^{-1})	1.2
k_e (min^{-1})	0.6160	k_{dest} (nM^{-1} min^{-1})	0.005
k_b (min^{-1} nM^{-1})	1.1690×10^{-5}	k_a (nM^{-1} min^{-1})	0.1
k_{mi} (min^{-1})	8.7234×10^{-5}	k_d (min^{-1})	0.001
k_{mo} (min^{-1})	5.4631	*Factor*	20
k_{dl} (min^{-1})	2.1061	k_{Wee1} (nM^{-1} min^{-1})	0.1
k_{dh} (min^{-1})	0.0120	k_{CDC25} (nM^{-1} min^{-1})	0.1
v_0	1.4778×10^{-5}	$\text{CDC2}_{\text{total}}$ (nM)	100
B_T (nM)	$1.3412\times10^{+5}$	$\text{APC}_{\text{total}}$ (nM)	50
V_{m1} (min^{-1} nM)	$9.6426\times10^{+2}$	$\text{Wee1}_{\text{total}}$ (nM)	15
V_{m2} (min^{-1} nM)	$2.40516\times10^{+2}$	$\text{CDC25}_{\text{total}}$ (nM)	15
k_{m1} (nM)	1.1400×10^{-3}	n_{APC}	5
k_{m2} (nM)	9.6828×10^{-3}	n_{Wee1}	4
k_{om} (min^{-1})	0.0096	n_{CDC25}	4
k_{cm} (min^{-1})	0.0180	n_{p21}	5
k_{oc} (min^{-1})	0.0096	EC50_{APC} (nM)	30
k_{cc} (min^{-1})	0.0180	$\text{EC50}_{\text{Wee1}}$ (nM)	25
v_1	0.3602	$\text{EC50}_{\text{CDC25}}$ (nM)	25
v_2	3.9732×10^{-1}	EC50_{p21} (nM)	30
v_3	1.0238×10^{-2}	k_{APCon} (min^{-1})	0.8
		k_{Wee1on} (min^{-1})	0.8
		k_{CDC25on} (min^{-1})	0.8
		k_{p21on} (min^{-1})	0.8
		k_{APCoff} (min^{-1})	0.08
		k_{Wee1off} (min^{-1})	0.08
		k_{CDC25off} (min^{-1})	0.08
		k_{CAK} (min^{-1})	0.8
		k_{PP2C} (min^{-1})	0.008

6.4 Structural identifiability analysis

6.4.1 Structural identifiability analysis of the drug kinetics model including the effect compartment

In this subsection, the Taylor series approach (see Section 2.4) is applied to the drug kinetics model (including the effect compartment γH2AX) described by Equations (6.1) with initial conditions given in Equation (6.2). The input in this particular experiment is a bolus injection of the anti-cancer drug TPT into the system, therefore, the corresponding amounts are included in the initial conditions $\mathbf{x}_0(\mathbf{p})$. The vector $\mathbf{p}$ comprising the two unknown parameters of the model is given by:

$$\mathbf{p} = (k_{\mathrm{TPT}}, k_{dam})^{\mathsf{T}}.$$

The unknown model parameters are positive, therefore, the set of feasible parameter vectors, Ω, comprises the vectors $(p_1, p_2)^{\mathsf{T}}$ such that $p_i > 0$ ($i = 1, 2$). The state space vector $\mathbf{x}(t, \mathbf{p})$ is given by:

$$\begin{aligned}\mathbf{x}(t,\ \mathbf{p}) = (&L_m(t,\ \mathbf{p}), H_m(t,\ \mathbf{p}), L_e(t,\ \mathbf{p}), H_e(t,\ \mathbf{p}), L_c(t,\ \mathbf{p}), H_c(t,\ \mathbf{p}),\\ &L_n(t,\ \mathbf{p}), \gamma H2AX(t,\ \mathbf{p}))^{\mathsf{T}}\end{aligned} \tag{6.16}$$

and the set $M(\mathbf{p})$ is $\mathbb{R}^8$. The initial condition for the state space representation of the model is given by:

$$\mathbf{x}_0(t,\ \mathbf{p}) = ((1+v_0)D, 0, 0, 0, 0, 0, 0, 0)^{\mathsf{T}}.$$

where D is the known dose and is therefore not included in the unknown parameter vector $\mathbf{p}$. The output function of the extended model is given by:

$$\mathbf{y}(t,\ \mathbf{p}) = \gamma H2AX(t,\ \mathbf{p}) \tag{6.17}$$

An arbitrary parameter vector is denoted by $\bar{\mathbf{p}}$ such that:

$$\bar{\mathbf{p}} = (\bar{k}_{\mathrm{TPT}}, \bar{k}_{dam})^{\mathsf{T}}$$

for which, $\mathrm{y}(t, \mathbf{p}) = \mathrm{y}(t, \bar{\mathbf{p}})$ for all $t \geq 0$. In order to perform the Taylor Series method (Section 2.3), successive coefficients of the Taylor series expansions of the components of $\mathrm{y}(t, \mathbf{p})$ and $\mathrm{y}(t, \bar{\mathbf{p}})$ are computed to determine the relationship between $\mathbf{p}$ and $\bar{\mathbf{p}}$. Similar to the structural identifiability analysis for the cell-based drug kinetics model in Chapter 4 and the cell cycle model in Chapter 5, the computer algebra system MATHEMATICA [206] was used to perform the following structural identifiability analysis of the extended cell-based drug kinetics model including the effect compartment (see Appendix A). As explained in Subsection 2.3.3, uniqueness of the coefficients in the Taylor series expansions of $y(t, \mathbf{p})$ entails the following, if $\bar{\mathbf{p}} \in \boldsymbol{\Omega}$ is such that $\mathrm{y}(t, \mathbf{p}) = \mathrm{y}(t, \bar{\mathbf{p}})$ for all $t \geq 0$, then, for each $k = 1, 2, 3, \ldots$

$$y^{(k)}(0,\mathbf{p}) = y^{(k)}(0,\bar{\mathbf{p}}) \text{ where, } y^{(0)}(0,\mathbf{p}) = y(0,\mathbf{p}). \tag{6.18}$$

For $k = 0, 1, 2, 3$ in Equation (6.18) no information is gained since each of these coefficients is 0, i.e., $y^{(k)}(0,\mathbf{p}) = 0$ for $k = 0, 1, 2, 3$. Now considering the 4th derivative ($k = 4$) yields:

$$y^{(4)}(0,\mathbf{p}) = \frac{B_T D k_{\mathrm{TPT}} k_b k_i k_{mi} \alpha}{v_2} \tag{6.19}$$

and substituting in Equation (6.18) gives:

$$\bar{k}_{\mathrm{TPT}} = k_{\mathrm{TPT}} . \tag{6.20}$$

The 5th derivative term ($k = 5$) gives:

$$\begin{aligned} y^{(4)}(0,\mathbf{p}) = & \frac{-B_T D k_{\mathrm{TPT}} k_b k_i k_{mi} (k_{dam} k_{m2} + B_T k_b k_{m2} + k_{dh} k_{m2} + k_{dl} k_{m2}) \alpha}{k_{m2} v_2} \\ & \frac{-B_T D k_{\mathrm{TPT}} k_b k_i k_{mi} (k_e k_{m2} + k_i k_{m2} + k_{m2} k_{mi}) \alpha}{k_{m2} v_2} \\ & \frac{-B_T D k_{\mathrm{TPT}} k_b k_i k_{mi} (k_{m2} k_{mo} + k_{m2} k_{oc} + 2 k_{m2} k_{om} + V_{m2}) \alpha}{k_{m2} v_2} \end{aligned} \tag{6.21}$$

and using the relation in Equation (6.20) yields:

$$\bar{k}_{dam} = k_{dam} . \tag{6.22}$$

Therefore, for Equation (6.18) to be satisfied, it is necessary that:

$$\bar{k}_{\mathrm{TPT}} = k_{\mathrm{TPT}} \text{ and } \bar{k}_{dam} = k_{dam} , \tag{6.23}$$

that is, $\mathbf{p} = \bar{\mathbf{p}}$ (for all $k \geq 0$). This is true for generic $\mathbf{p} \in \Omega$, thus the model is structurally globally identifiable, that is all of the unknown model parameters in **p** are uniquely determined by the output structure corresponding to the proposed experiments to collect data for the purpose of parameter estimation. In Section 6.5, the unknown model parameters are estimated from high-content imaging data for the identification and counting of γH2AX. The structural identifiability analysis was performed in this section to approach the estimation of the unknown model parameters with more confidence.

6.4.2 Structural identifiability analysis of the full drug kinetics/dynamics model

In this subsection, the Taylor series approach (see Subsection 2.3.3) is applied to the coupled drug kinetics/cell cycle response (dynamics) model at time $t = t_{G1}$ (the start of a new cell cycle, see Sections Chapter 5) described by Equations (6.1) and Equations (6.5)-(6.14). For this study, the output function, $y(t, \mathbf{p})$, of the coupled drug kinetics/cell cycle response model is taken to be a linear function of the active Cyclin B1 concentration ($[CDC2CycB_{TP}]$), that is:

$$y(t, \mathbf{p}) = \alpha CDC2CycB_{TP}(t, \mathbf{p}) + \beta \tag{6.24}$$

where, α and β are unknown observational parameters estimated from the experimental data (see Section 5.4). The state space model describing how the drug perturbs the cell cycle process is an uncontrolled non-linear system.

The vector **p** comprising the five unknown parameters of the model is given by:

$$\mathbf{p} = (\alpha, \beta, C_2, k_{p21off}, p21_{total})^{\top}.$$

The unknown model parameters represent observational parameters (α and β), a scaling factor (C_2), a rate constant (k_{p21off}) and a concentration ($p21_{total}$), and are therefore positive. This means that the set of feasible parameter vectors, Ω, comprises the vectors $(p_1, \ldots, p_5)^{\top}$ such that $p_i > 0$ $(1 \leq i \leq 5)$. The state space vector $\mathbf{x}(t, \mathbf{p})$ is given by:

$$\begin{aligned}\mathbf{x}(t,\ \mathbf{p}) = (&L_m(t,\ \mathbf{p}), H_m(t,\ \mathbf{p}), L_e(t,\ \mathbf{p}), H_e(t,\ \mathbf{p}), L_c(t,\ \mathbf{p}), H_c(t,\ \mathbf{p}),\\ &L_n(t,\ \mathbf{p}), \gamma H2AX(t,\ \mathbf{p}), \mathrm{CycB}(t,\ \mathbf{p}), \mathrm{CDC2CycB}(t,\ \mathbf{p}),\\ &\mathrm{CDC2CycB_{YP}}(t,\ \mathbf{p}), \mathrm{CDC2CycB_{YPTP}}(t,\ \mathbf{p}), \mathrm{CDC2CycB_{TP}}(t,\ \mathbf{p}),\\ &\mathrm{CDC2}(t,\ \mathbf{p}), \mathrm{APC_{active}}(t,\ \mathbf{p}), \mathrm{Wee1_{active}}(t,\ \mathbf{p}),\\ &\mathrm{CDC25_{active}}(t,\ \mathbf{p}), \mathrm{p21_{active}}(t,\ \mathbf{p}))^{\mathsf{T}}\end{aligned} \tag{6.25}$$

and the set $M(\mathbf{p})$ is $\mathbb{R}^{18}$. The values for the state variables of the model at $t = t_{G1}$ are given by:

$$\begin{aligned}\mathbf{x}(t_{G1},\ \mathbf{p}) = (&a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, 0.141396C_2,\\ &= 0.992654C_2, 1.4763C_2, 11.41023C_2, 1.64121C_2,\\ &\ \ 84.479606C_2, 3.8539C_2, 15C_2, 1.54009C_2, 3.8539C_2)^{\mathsf{T}}\end{aligned} \tag{6.26}$$

where a_i (i = 1, …, 8) are the values of the variables in the model described by Equation (6.1) (i.e., the drug kinetics model including the effect compartment) which depend on the time point (t_{G1}) at which a new cell cycle starts and on the value of the dose D (treatment regimen). Determining the values of a_i (i = 1, …, 8) is explained in Section 6.5. In this model (i.e., the coupled drug kinetics/dynamics model), the values for the state variables at time $t = t_{G1}$ are actually dependent on components of the vector of unknown parameters **p**.

An arbitrary parameter vector is denoted by $\bar{\mathbf{p}}$ such that:

$$\bar{\mathbf{p}} = (\bar{\alpha}, \bar{\beta}, \bar{C}_2, \bar{k}_{\mathrm{p21off}}, \bar{\mathrm{p}}21_{\mathrm{total}})^{\mathsf{T}}$$

for which, $\mathrm{y}(t,\ \mathbf{p}) = \mathrm{y}(t,\ \bar{\mathbf{p}})$ for all $t \geq 0$. Similar to the structural identifiability analysis for the cell-based kinetics model (Chapter 4) and the extended cell cycle model (Chapter 5), the computer algebra system MATHEMATICA [206] was used to perform the following structural identifiability analysis of the extended cell cycle model (see Appendix A). As explained in Subsection 2.3.3, uniqueness of the coefficients in the Taylor series expansions of $y(t,\ \mathbf{p})$ entails the following, if $\bar{\mathbf{p}} \in \boldsymbol{\Omega}$ is such that $\mathrm{y}(t, \mathbf{p}) = \mathrm{y}(t,\ \bar{\mathbf{p}})$ for all $t \geq 0$, then, for each k = 1, 2, 3, …

$$y^{(k)}(t_{G1}, \mathbf{p}) = y^{(k)}(t_{G1}, \bar{\mathbf{p}}) \text{ where, } y^{(0)}(t_{G1}, \mathbf{p}) = y(t_{G1}, \mathbf{p}). \tag{6.27}$$

Setting $k = 0$ in Equation (6.27) yields the first Taylor series coefficient given by:

$$y^{(0)}(t_{G1}, \mathbf{p}) = 1.64121C_2\alpha + \beta \tag{6.28}$$

and substituting in Equation (6.27) gives:

$$\bar{\beta} = 1.64121 C_2 \alpha - 1.64121 \bar{C}_2 \bar{\alpha} + \beta. \tag{6.29}$$

The 1st derivative term ($k = 1$) gives:

$$y^{(1)}(t_{G1}, \mathbf{p}) = 2.52518 C_2^{\,2} \alpha \tag{6.30}$$

and using the relations in Equation (6.29) yields:

$$\bar{\alpha} = \frac{C_2^{\,2} \alpha}{\bar{C}_2^{\,2}}. \tag{6.31}$$

Considering the relations (Equations (6.29) and (6.31)) between parameters in $\mathbf{p}$ and $\bar{\mathbf{p}}$, the following equations must hold to satisfy Equation (6.27) for $k = 0, 1$:

$$\bar{\alpha} = \frac{C_2^{\,2} \alpha}{\bar{C}_2^{\,2}}, \text{ and } \bar{\beta} = 1.64121 C_2 \alpha - 1.64121 \frac{C_2^{\,2} \alpha}{\bar{C}_2} + \beta. \tag{6.32}$$

The 2nd derivative term ($k = 2$) gives:

$$y^{(2)}(t_{G1}, \mathbf{p}) = (0.107836 - 0.388901 C_2) C_2^{\,2} \alpha \tag{6.33}$$

and using the relations in Equations (6.32) and (6.33) yields:

$$\bar{C}_2 = C_2. \tag{6.34}$$

Considering the relations, in Equations (6.32) and (6.34), between parameters in $\mathbf{p}$ and $\bar{\mathbf{p}}$, the following equations must hold to satisfy Equation (6.27) for $k = 0, 1, 2$:

$$\bar{C}_2 = C_2, \ \bar{\alpha} = \alpha, \text{ and } \bar{\beta} = \beta. \tag{6.35}$$

The 3rd derivative term ($k = 3$) gives:

$$y^{(3)}(t_{G1}, \mathbf{p}) = 0.0598942(-0.0627616 + C_2) C_2^{\,2} (4.45049 + C_2) \alpha \tag{6.36}$$

and substituting in Equation (6.27) yields no information since none of the remaining unknown model parameters are present in the 4th coefficient of Taylor series expansion. The 4th derivative term ($k = 4$) gives:

$$\begin{aligned} y^{(4)}(t_{G1}, \mathbf{p}) = {} & 0.00252426 C_2^{\,2} \alpha - 0.0345519 C_2^{\,3} \alpha - 0.171215 C_2^{\,4} \alpha \\ & - 0.00922425 C_2^{\,5} \alpha \end{aligned} \tag{6.37}$$

and substituting in Equation (6.27) yields no information since none of the remaining unknown model parameters are present in the 5th coefficient of Taylor series expansion. The 5th derivative term ($k = 5$) gives:

$$\begin{aligned} y^{(5)}(t_{\mathrm{G1}}, \mathbf{p}) = {} & C_2{}^2\alpha(-0.000310113 + 2.6964C_2 - 2.50228C_2{}^3 \\ & +0.00142062C_2{}^4 + C_2{}^2(0.0542277 - 5.16716\times10^{-10}\,\mathrm{p21_{total}})) \\ & + \frac{0.00507981\, k_{\mathrm{p21off}} C_2{}^4\alpha}{(1 + a_8 k_{dam})} \end{aligned} \tag{6.38}$$

and using the relations in Equations (6.35) yields:

$$\begin{aligned} \bar{k}_{\mathrm{p21off}} = {} & k_{\mathrm{p21off}} - (1.01719\times10^{-7} + 1.01719\times10^{-7} a_8 k_{dam})\mathrm{p21_{total}} \\ & + (1.01719\times10^{-7} + 1.01719\times10^{-7} a_8 k_{dam})\bar{\mathrm{p}}21_{\mathrm{total}} \end{aligned}. \tag{6.39}$$

Finally, calculating the seventh coefficient of Taylor series expansion ($k = 6$), and using the relations in Equations (6.37) and (6.41) gives:

$$\bar{\mathrm{p}}21_{\mathrm{total}} = \mathrm{p21_{total}}. \tag{6.40}$$

Therefore, the following relations must hold in order to satisfy Equation (6.27) for all $k = 0, \ldots, 6$:

$$\bar{C}_2 = C_2,\ \bar{k}_{\mathrm{p21off}} = k_{\mathrm{p21off}},\ \bar{\mathrm{p}}21_{\mathrm{total}} = \mathrm{p21_{total}},\ \bar{\alpha} = \alpha \text{ and } \bar{\beta} = \beta. \tag{6.41}$$

Equation (6.27) is therefore satisfied for all $k \geq 0$ with $\mathbf{p} = \bar{\mathbf{p}}$. This is true for generic $\mathbf{p} \in \Omega$, thus the model is structurally globally identifiable, that is all of the model parameters are uniquely determined by the output structure corresponding to the proposed experiments to collect data for the purpose of parameter estimation. In Section 6.5.2, the unknown model parameters were estimated from time-lapse microscopy data (Cyclin B1-GFP fluorescence data) at the start of a new cell cycle (i.e., at $t = t_{\mathrm{G1}}$), in the presence of the anti-cancer agent TPT. The structural identifiability analysis was performed in this section to approach the estimation of the unknown model parameters with greater confidence.

6.5 Parameter estimation of the coupled drug kinetics/cell cycle model

6.5.1 Parameter estimation before the *washout* event

In this subsection, the DNA damage data (the number of γH2AX foci in single nuclei) for live human osteosarcoma cells (U-2 OS cell line) treated with $1\times10^{+3}$ and $10\times10^{+3}$ nM TPT were used to estimate the unknown model parameters, namely, k_{TPT} and k_{dam} (for the two treatments regimens) prior to the *washout* event at time $t_w = 60$ min, see Section 6.3. Figure 6.4 illustrates distinctly different patterns of response for these cells to the different treatment regimens. The cells were treated with a bolus of TPT and after 60 min of administration, the number of γH2AX foci per nucleus are quantified for individual cells (1 data point at $t_w = 60$ min). The average DNA damage (i.e., the dsDNA breaks) induced by TPT (see Table 6.2), increases with the value of dose administered as shown in Figure 6.4.

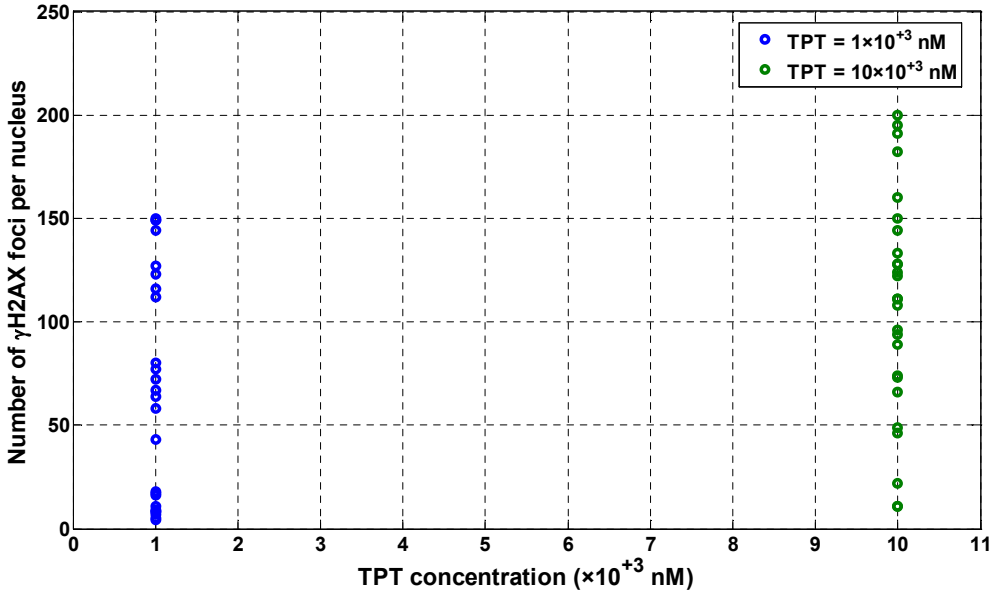


Figure 6.4: The number of γH2AX foci in single nuclei for live human osteosarcoma cells treated for one hour with $1\times10^{+3}$ nM (blue) and $10\times10^{+3}$ nM (green) TPT.

Using the estimated drug kinetics model parameters in Table 6.1, the FACSIMILE software (Appendix B) package (MCPA Software, U.K.) was used to estimate the unknown model

parameters k_{TPT} and k_{dam} for the two treatment regimens (Table 6.2). Data from 25 cells (for those treated with $1\times10^{+3}$ nM TPT) and 26 cells (for those treated with $10\times10^{+3}$ nM TPT) were averaged as presented in Table 6.2, and the postulated model (Equations (6.1)) fitted to these averaged data. The estimated parameter values and the predicted number of γH2AX foci per nucleus for each treatment regimen after 60 min are presented in Table 6.2. The predicted count of γH2AX foci per nucleus show that there is exact reproduction of experimental data (in both experiments) by the simulated output from the model taking parameter values from Tables 6.1 and 6.2.

Table 6.2: Parameter values for the model described by Equations (6.1), estimated using high-content imaging.

TPT = $1\times10^{+3}$ nM		**TPT = $10\times10^{+3}$ nM**	
Parameter	**Value**	**Parameter**	**Value**
k_{TPT} (nM min)$^{-1}$	2.1885×10^{-3}	k_{TPT} (nM min)$^{-1}$	1.83468×10^{-4}
k_{dam} (min^{-1})	1.43934×10^{-2}	k_{dam} (min^{-1})	4.0782×10^{-3}
Average γH2AX foci per nucleus	**Predicted γH2AX foci per nucleus**	**Average γH2AX foci per nucleus**	**Predicted γH2AX foci per nucleus**
65	65	108	108

The value of the parameter k_{TPT} decreases with increasing dose. In addition, the value of this parameter (i.e., k_{TPT}) in both experiments is very small in comparison to the other parameters relating to compartment L_n (drug bound in the nucleus). Thus the estimated value for the parameter k_{TPT} is very small in comparison with any of k_{dl}, k_{dh}, in addition to the product of B_T and k_b (see Table 6.1). Therefore, extending the drug kinetics model presented in Chapter 4 (Figure 4.9) by introducing the effect compartment γH2AX (see Figures 6.3 and 6.4) has a *negligible* effect on the drug kinetics model. Moreover, the value of the parameter k_{dam} decreases with increasing dosage (D), and consequently the intensity of response to the drug (i.e., the DNA damage induced by the drug) increases.

The drug kinetics model described by Equations (6.1) was simulated for 60 min, with parameters taking values from Tables 6.1 and 6.2 and treatment regimens $D = 1\times10^{+3}$ and $10\times10^{+3}$ nM TPT as shown in Figures 6.5-6.8. The average number of γH2AX foci per nucleus increases as the concentration of bound drug to the nucleus (L_n) increases (Figures 6.6 and 6.8). Moreover, when cells are treated with $10\times10^{+3}$ nM TPT, the concentration L_n

reaches a maximum value of $1.2457\times10^{+4}$ nM at time $t = 13.5$ mins (Figure 6.7), and then starts to gradually decrease. However, the number of γH2AX foci continues to increase; this indicates that DNA damage that generates the formation of dsDNA breaks is still increasing and the DNA repair process does not start immediately once the concentration of the drug bound to its DNA target decreases.

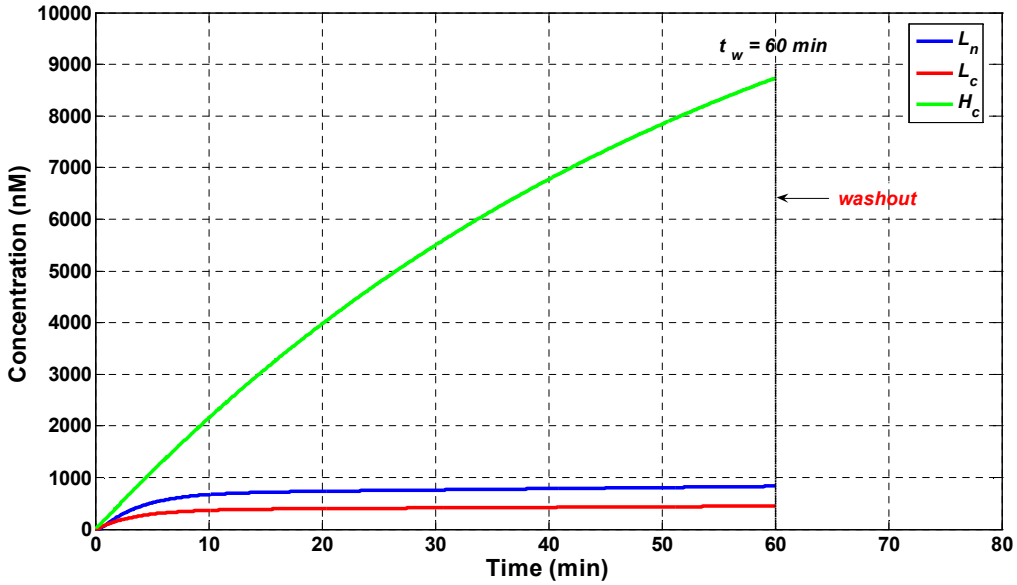


Figure 6.5: Simulated output for L_n (blue), L_c (red) and H_c (green) from model Equation (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 1\times10^{+3}$ nM TPT for 60 min (until the *washout* event).

The predicted concentration for the inactive form of the drug (TPT_H) in the cytoplasm (H_c) diverges over time (in both experiments) as shown in Figures 6.5 and 6.7. The pH level at which the experiments were undertaken was 7.2; the results obtained are consistent with previous studies [98, 99] that TPT_L in the cytoplasm (L_c) is unstable at physiological pH and undergoes reversible hydrolysis to TPT_H (H_c). The enzyme ALDH is also responsible for the irreversible hydrolysis (inactivation) of TPT_L to TPT_H (in the cytoplasm) which brings about the increase in the ratio of TPT_H to TPT_L in the cytoplasm during the first 60 min (see Chapter 4). In addition, the dissociation of L_n to H_c increases the concentration of H_c. For both treatment regimens, the predicted concentrations of the active form of the drug (TPT_L) in the cytoplasm (L_c) and in the nucleus (L_n) settle to a steady state, as shown in Figures 6.5 and 6.7.

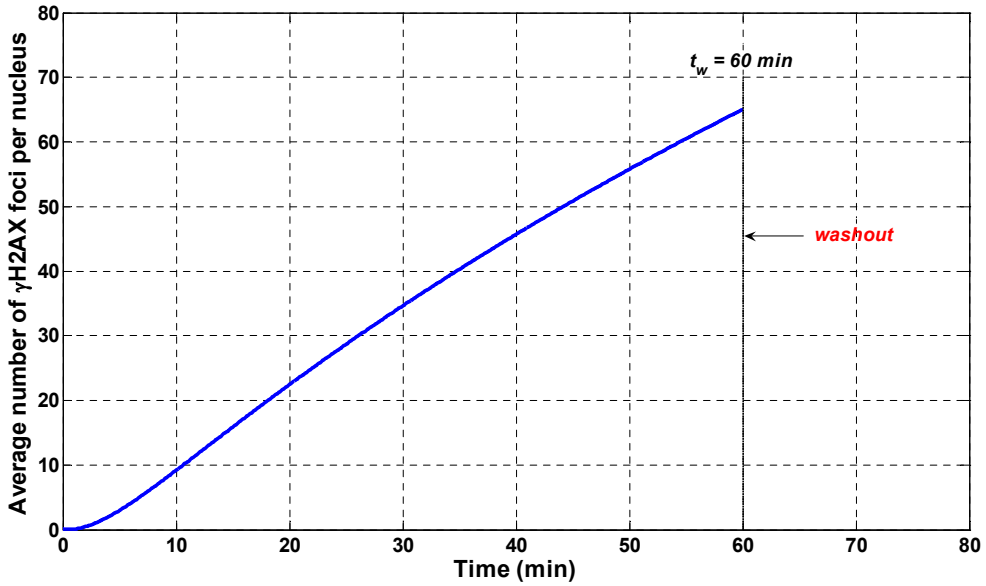


Figure 6.6: Simulated output for the average number of γH2AX foci per nucleus from the model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 1\times10^{+3}$ nM TPT for 60 min (until the *washout* event).

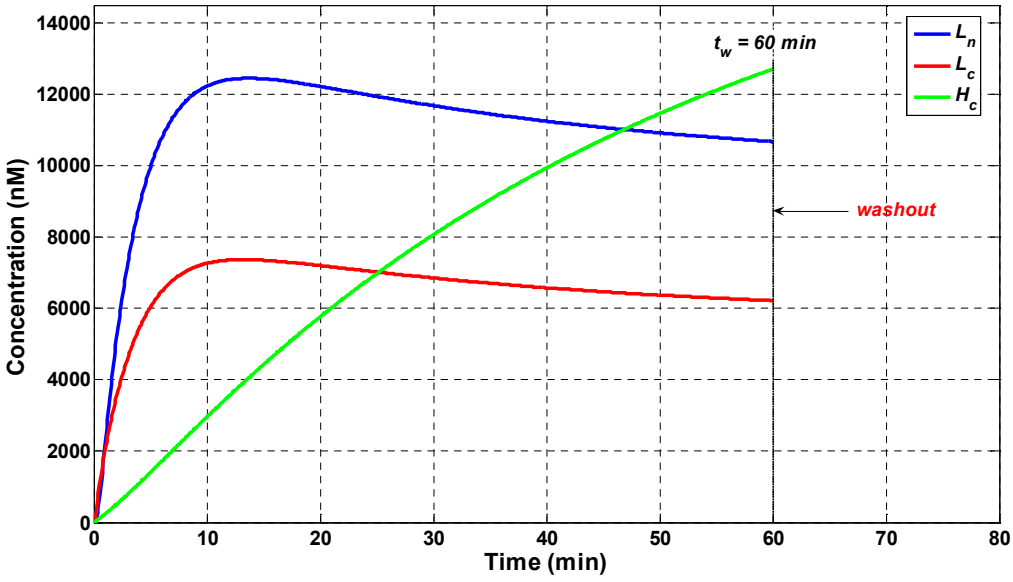


Figure 6.7: Simulated output for L_n (blue), L_c (red) and H_c (green) from model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 10\times10^{+3}$ nM TPT for 60 min (until the *washout* event).

Following the *washout* event at time $t_w = 60$ min, the variables relating to the drug kinetics model (including the effect compartment) that are unaffected by this event for $D = 1\times10^{+3}$ nM are given by (see Figures 6.5 and 6.6):

$$L_c(60) = 0.45\times10^{+3} \text{ nM},\ H_c(60) = 8.73\times10^{+3} \text{ nM},$$
$$L_n(60) = 0.833\times10^{+3} \text{ nM and } \gamma H2AX(60) = 65 \tag{6.42}$$

and the corresponding values for these variables at time $t_w = 60$ min when $D = 10\times10^{+3}$ nM are given by (see Figures 6.7 and 6.8):

$$L_c(60) = 6.22\times10^{+3} \text{ nM},\ H_c(60) = 1.27\times10^{+4} \text{ nM},$$
$$L_n(60) = 1.07\times10^{+4} \text{ nM and } \gamma H2AX(60) = 108. \tag{6.43}$$

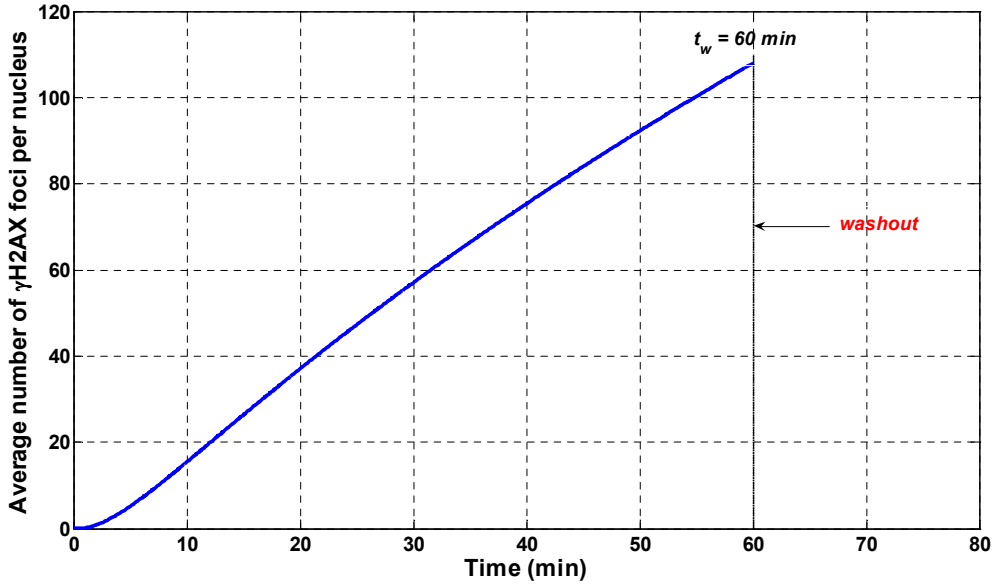


Figure 6.8: Simulated output for the average number of γH2AX foci per nucleus from the model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 10\times10^{+3}$ nM TPT for 60 min (until the *washout* event).

6.5.2 Parameter estimation of the coupled model

In this chapter, the coupled drug kinetics/dynamics model describes the response of single human osteosarcoma cells treated with $1\times10^{+3}$ nM and $10\times10^{+3}$ nM TPT. Similar to the approach used for the control (untreated) cell in Chapter 5, the unknown model parameters C_2, k_{p21off}, $\text{p21}_{\text{total}}$, α and β (as in Section 6.4) were estimated using Cyclin B1-GFP data (see Section 6.2) for one complete cell cycle event of a newborn cell starting at time $t = t_{\text{G1}}$ (see Chapter 5) to study the effects of TPT on cell cycle progression. In addition, the unknown model parameters were estimated from averaged cytoplasmic Cyclin B1-GFP data (see Figure 6.9) from the two ROI's (see Section 6.2). Moreover, due to the artefacts of morphological changes (rounding up of the cell before cell division) during the M phase of the cell cycle, which result in a false peak (Figure 6.9) in the data, the parameter

estimation was performed starting from time $t = t_{G1}$ (G_1 phase of a new cell cycle) for one complete cell cycle excluding the experimental time point corresponding to the false peak (see Section 5.2).

In this subsection, the FACSIMILE software package (MCPA Software, U.K.) was used to fit the coupled drug kinetics/cell cycle response model to Cyclin B1-GFP fluorescence data. The unknown parameters of the coupled model are the observational parameters (α and β), the concentration scaling factor C_2 (specific to cell type and cellular conditions, see Chapter 5), total concentration of the CKI $p21^{CIP1/WAF1}$ ($[p21_{total}]$) and the rate constant for turning $p21^{CIP1/WAF1}$ off (k_{p21off}). Live cells are tracked in time and therefore the unknown model parameters are permitted to change (if appropriate) at cytokinesis, however, they remain constant during an individual cycle. The parameter values in Table 6.1 are specific for all cell types and cellular conditions (internal and external) that influence the cell cycle progression. These parameters are also permitted to change implicitly (if appropriate) at cell division since they are scaled by C_1 or C_2. However, they remain constant during an individual cell cycle.

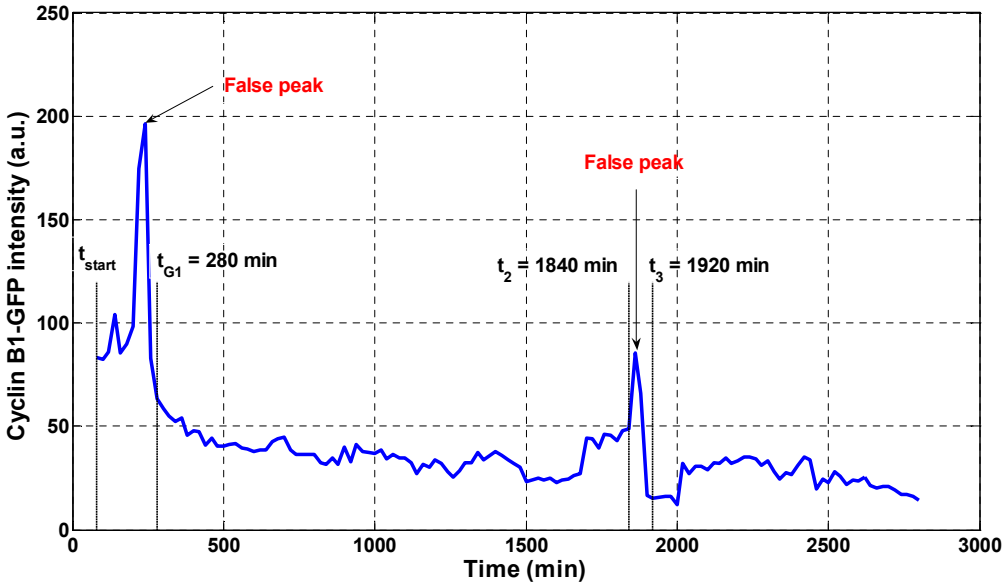


Figure 6.9: Average cytoplasmic Cyclin B1-GFP fluorescence intensity extracted from a typical G_2 cell treated with $1\times10^{+3}$ nM TPT.

Figure 6.9 represents average cytoplasmic Cyclin B1-GFP intensity profile from a G_2 cell responding to a dose $D = 1\times10^{+3}$ nM of TPT. The high-resolution fluorescence tracking

begins (i.e. the first image is captured) 20 min after the *washout* event, that is at time $t_{start} = 80$ min (see Figure 6.9). In this particular experiment, the progenitor cell divides into two daughter cells (first cell division) at time $t_B = 260$ min (see Chapter 5) and the complete newborn cell appears at time $t_{G1} = 280$ min (the time when a new cycle starts from the G_1 phase). The values for the state variables relating to the cell cycle response component of the coupled model at time t_{G1} are given in Equation 5.16 (see Chapter 5). In order to identify the values for the state variables of the full model (coupled drug kinetics/cell cycle response model), the drug kinetics model (including the effect compartment) was simulated to time $t_{G1} = 280$ min (i.e., 220 min following the *washout* event) as shown in Figures 6.10-6.12.

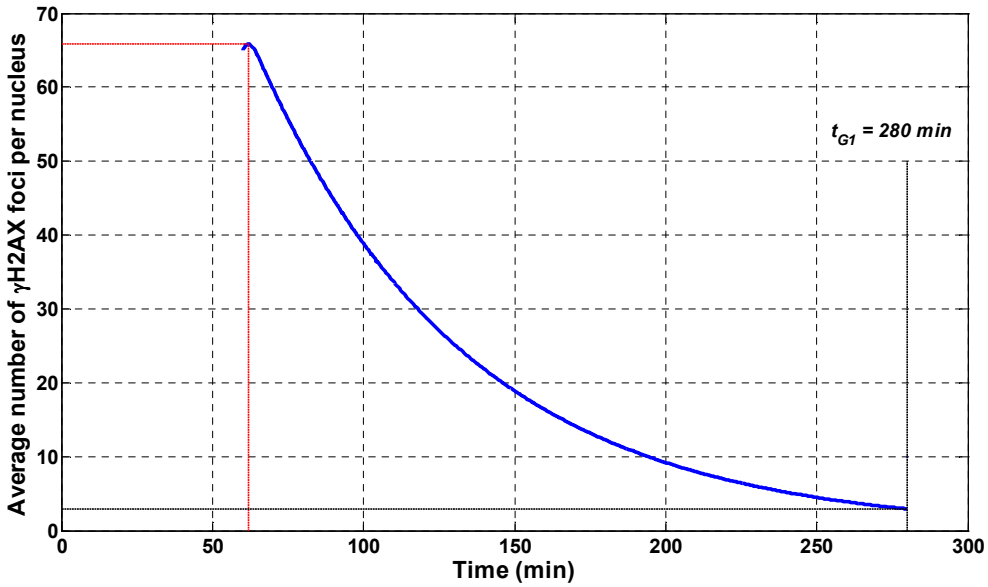


Figure 6.10: Simulated output for average number of γH2AX foci per nucleus from the model Equation (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 1\times10^{+3}$ nM following the *washout* event until time $t = t_{G1}$.

The average number of γH2AX foci per nucleus slightly increases and reaches the maximum number of γH2AX foci per nucleus at around time $t = 62$ min. Then, the number of γH2AX foci starts to exponentially decrease (see Figure 6.10), that is when the cell starts the DNA repair mechanism. The concentration of TPT_H in the cytoplasm (H_c) also increases following the *washout* event (Figure 6.11) and reaches a maximum concentration of $9.0821\times10^{+3}$ nM at time $t = 65.3333$ min as a result of the reversible hydrolysis of L_c to H_c, the irreversible hydrolysis of L_c to H_c (due to the catalytic reaction of the enzyme ALDH) and the dissociation of TPT_L bound to DNA (i.e., L_n) to H_c. Then, H_c decays to

zero (i.e., 0 nM, see Figure 6.16) over time as shown in Figure 6.11. The concentrations of both TPT_L in the cytoplasm (L_c) and nucleus (L_n) start rapidly decreasing after the *washout* event (Figure 6.12). Therefore, the drug in the cellular compartments flows to the extracellular region by simple diffusion (only TPT_L) or by active pumping (only TPT_H through the BCRP/ABCG2 transporter). These results match expected biological behaviour.

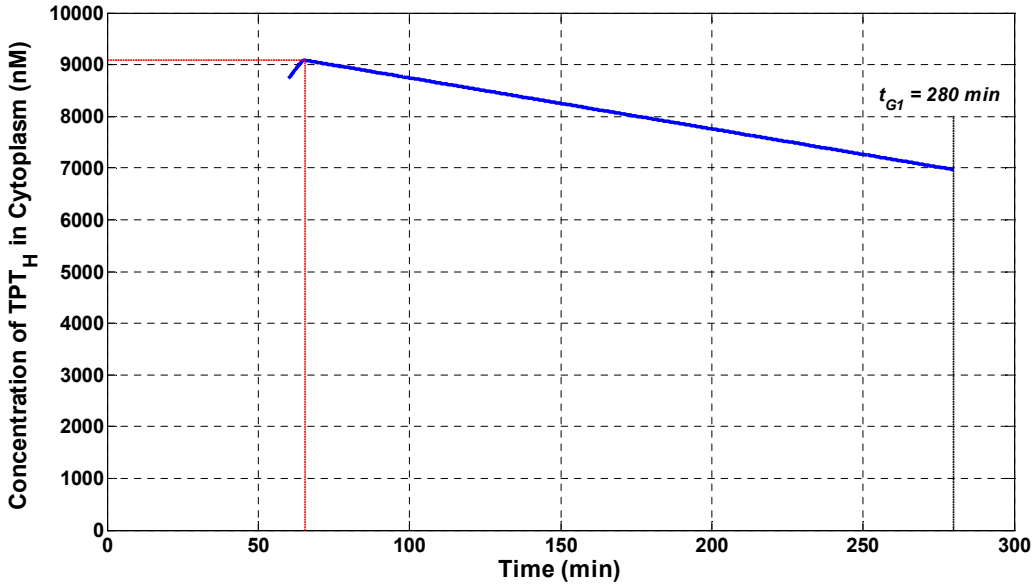


Figure 6.11: Simulated output for TPT_H in the cytoplasm (H_c) from model Equation (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 1\times10^{+3}$ nM TPT following the *washout* event until time $t = t_{G1}$.

The values for the state variables of the full drug kinetics/cell cycle response model (dynamics) at time $t_{G1} = 280$ min when cells are treated with $1\times10^{+3}$ nM TPT are given by:

$$\begin{aligned}
[\text{CycB}](280) &= 0.141396C_2, [\text{CDC2CycB}](280) = 0.992654C_2,\\
[\text{CDC2CycB}_{\text{YP}}](280) &= 1.4763C_2, [\text{CDC2CycB}_{\text{YPTP}}](280) = 11.41023C_2,\\
[\text{CDC2CycB}_{\text{TP}}](280) &= 1.64121C_2, [\text{CDC2}](280) = 84.479606C_2,\\
[\text{APC}_{\text{active}}](280) &= 3.8539C_2, [\text{Wee1}_{\text{active}}](280) = 15C_2,\\
[\text{CDC25}_{\text{active}}](280) &= 1.54009C_2, [\text{p21}_{\text{active}}](280) = 3.8539C_2,\\
L_m(280) &= 0.0584,\ H_m(280) = 0.0457,\ L_e(280) = 0.0659,\\
H_e(280) &= 5.0496,\ L_c(280) = 0.0106,\ H_c(280) = 6.97\times10^{+3},\\
L_n(280) &= 0.0197 \text{ and } \gamma H2AX(280) = 2.91
\end{aligned} \tag{6.44}$$

where, the state variable $\gamma H2AX$ is dimensionless and the rest of the state variables are measured in nM.

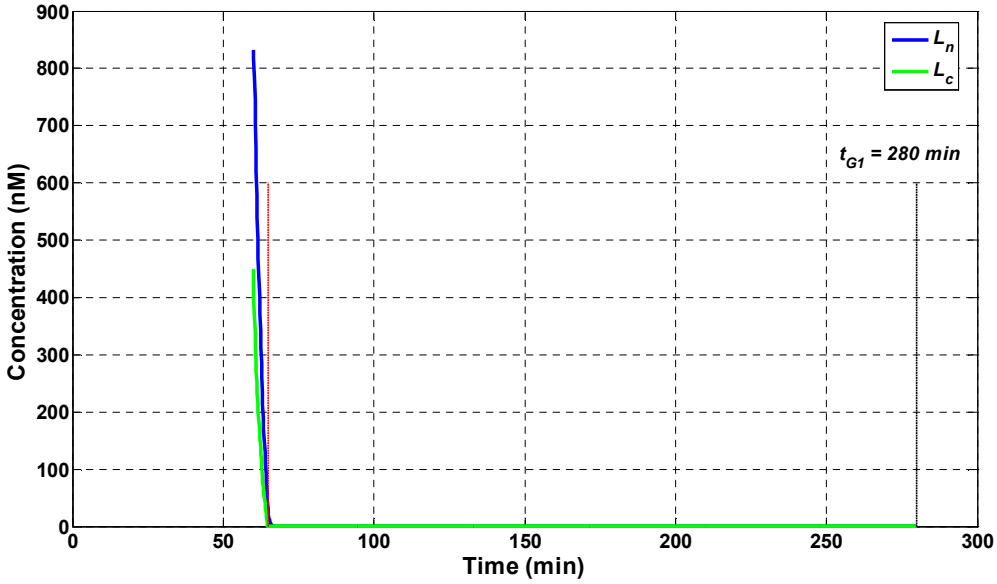


Figure 6.12: Simulated output for L_n (blue) and L_c (green) from model Equation (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 1\times10^{+3}$ nM TPT following the *washout* event until time $t = t_{G1}$.

In the track shown in Figure 6.9, the false peak is located at time $t = 1860$ min and the cell divides again (second cell division) at time $t = 1900$ min, therefore, a complete newborn cell appears at time $t_3 = 1920$ min (the time when a new cycle starts from the G_1 phase). The parameter estimation (for the cell treated with $1\times10^{+3}$ nM TPT) was performed starting from time $t_{G1} = 280$ min until time $t_3 = 1920$ min (one complete cell cycle) excluding the experimental data point which includes the false peak. Using Equation (5.35), the scaling factor C_1 is given by (see Section 5.4):

$$C_1 = \frac{\mathrm{IMT}_a}{\mathrm{IMT}_b}. \tag{6.45}$$

where, IMT_a is the inter-mitotic time in the extended model before scaling and IMT_b is the inter-mitotic time in the experimental data. IMT_a is a constant and is equal to 138.5 min (see Figure 5.19, Chapter 5) and for this particular track (Figure 6.9) IMT_b = 1640 min (1920 min – 280 min), and therefore, $C_1 = 0.0845$ (see Table 6.3).

The fitted parameter values, and their confidence levels are presented in Table 6.3, where it is seen that all parameters are well-determined by the Cyclin B1-GFP data. Additionally, three of the estimated parameters, namely, C_2, α and β, have low SDLN values corresponding to high confidence in their values. However, the estimated parameter values

for k_{p21off} and $p21_{total}$ have high SDLN values due to the low signal-to-noise ratio in the experimental data. Moreover, the value of $CORI_{mean} = 0.62857$ indicates that the residuals are acceptable and evenly distributed above and below the simulated output of the model and therefore the resulting fit is good (in terms of the residuals).

Table 6.3: Best parameter estimates for the coupled drug kinetics/cell cycle response model Equation (6.1)-(6.14) and state values at $t = t_{G1}$ in Equation (6.44), estimated using live-cell Cyclin B1-GFP data for a cell treated with $1\times10^{+3}$ nM.

Parameter				Value
C_1				0.0845
Parameter	**Value**	**SDLN**	**5%**	**95%**
α (a.u. nM^{-1})	2	0.3483	1.13	3.55
C_2	9×10^{-1}	0.0275	8.60×10^{-1}	9.42×10^{-1}
β (a.u.)	$2.83\times10^{+1}$	0.1214	$2.32\times10^{+1}$	$3.46\times10^{+1}$
k_{p21off} (min^{-1})	7.20×10^{-2}	1.4102	7.08×10^{-3}	7.32×10^{-1}
$p21_{total}$ (nM)	$1.85\times10^{+2}$	1.7707	$1.00\times10^{+1}$	$3.40\times10^{+3}$
RSS = $7.1229\times10^{+1}$				

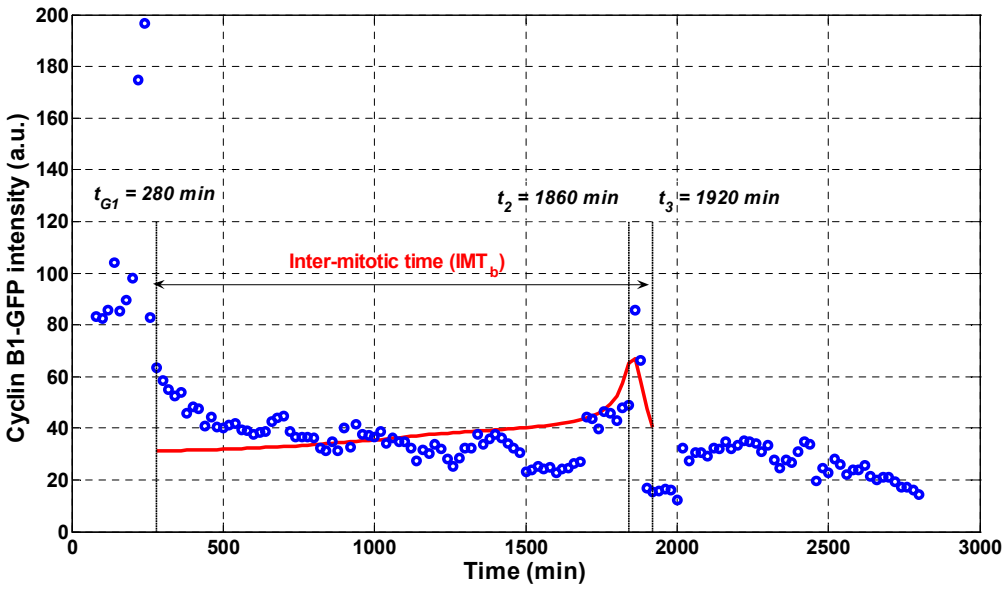


Figure 6.13: Comparison of model prediction (solid red line) and Cyclin B1-GFP data (blue circles) in the presence of $1\times10^{+3}$ nM TPT. An artefact of morphological changes during the final stages of mitosis is the over-expression of Cyclin B1 fluorescence giving rise to false peaks in the experimental data.

The plots shown in Figure 6.13 demonstrate a typical trace for the growth of a cell treated with $1\times10^{+3}$ nM TPT. The Cyclin B1-GFP data are denoted by the blue circles and the simulated data by the solid red curve. The model fit in Figure 6.13 does not show that there is a close reproduction of Cyclin B1-GFP data by simulated output from the coupled model

(with parameters taking values from Tables 6.1-6.3) due to the noise in the data. However, the simulated output from the model detects the exact location of the peak (spike) in the Cyclin B1-GFP data and reproduces the general phenotypic (qualitative) behaviour (i.e., position in the cell cycle) of the experimental data. Similar to the extended cell cycle model presented in Chapter 5, the impact of the false peak (see Figure 6.9) in the experimental data, due to morphological changes (rounding up) during the M phase of the cell cycle, is minimised by the coupled drug kinetics/cell cycle response model.

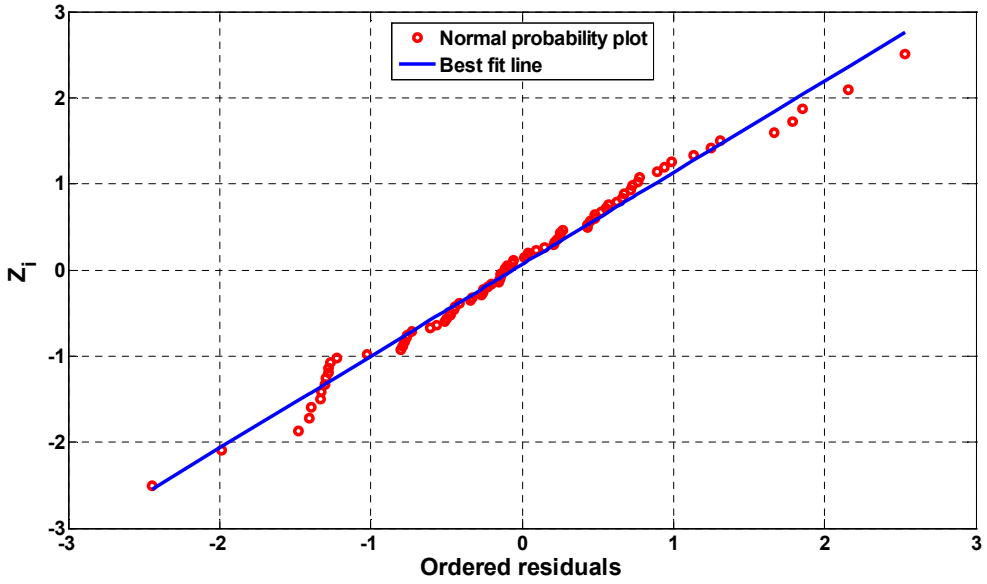


Figure 6.14: A normal probability plot for the residuals in the weighted non-linear least-squares fit of the coupled drug kinetics/cell cycle response model to Cyclin B1-GFP data extracted from typical G_2 cell treated with $1\times10^{+3}$ nM TPT. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

Figure 6.14 shows the normal probability plot for the weighted residuals for the fit in Figure 6.13 of the coupled drug kinetics/cell cycle response model to Cyclin B1-GFP data of a cell treated with $1\times10^{+3}$ nM TPT. The residuals are plotted in ascending order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale. The figure is consistent with the assumption of normally distributed errors (with approximately zero mean and standard deviation of 1), since the graph is approximately linear, which is an indicator for the appropriateness of the standard error σ used for each observation y.

The estimated correlation matrix for the well-determined parameters is given in Table 6.4. From this table the parameter pairings that exhibit the highest correlation are: $p21_{total}$ and k_{p21off} (0.861) and α and β (-0.916). The first pair represent the total concentration of CKI $p21^{CIP1/WAF1}$ and the rate at which this CKI is turned off. The high correlation of the logarithm of these parameters indicates that it is the product of these that is important for the fitting, i.e., the rate at which the inhibitor is activated. The second pair relates to the experimental observations. Accordingly, the high correlation of the logarithm of these parameters (α and β) indicates that it is the sum of these parameters that is important for the fitting.

Table 6.4: The estimated correlation matrix for the well-determined parameters for a cell treated with $1\times10^{+3}$ nM TPT.

	A	C_2	β	k_{p21off}	$p21_{total}$
α	1	-0.416	-0.916	0.601	0.795
C_2	-0.416	1	0.441	-0.946	-0.662
β	-0.916	0.441	1	-0.593	-0.729
k_{p21off}	0.601	-0.946	-0.593	1	0.861
$p21_{total}$	0.795	-0.662	-0.729	0.861	1

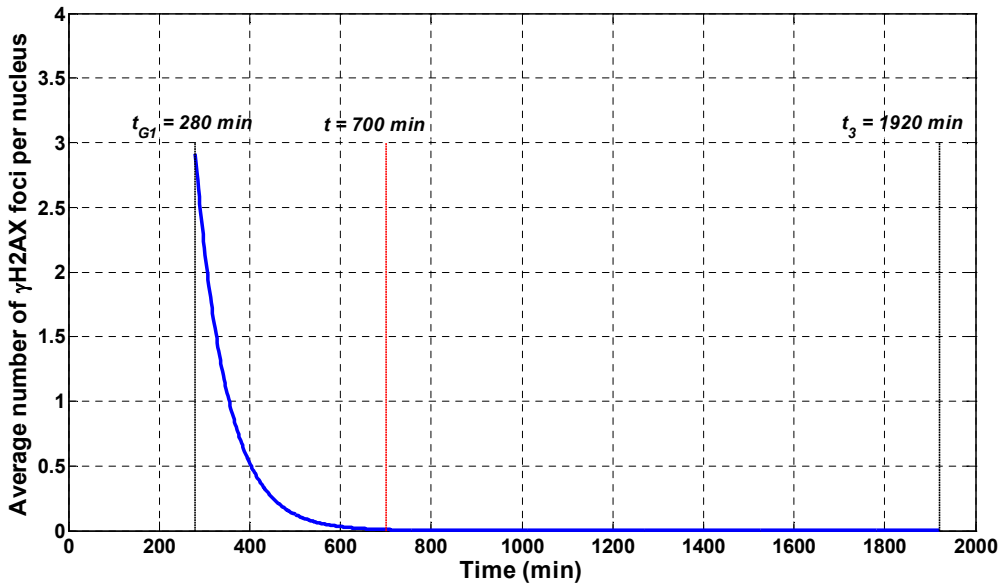


Figure 6.15: Simulated output for the average number of γH2AX foci per nucleus from the model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 1\times10^{+3}$ nM during one complete cell cycle.

Figure 6.15 represents the simulated output for the average number of γH2AX foci per nucleus during a complete cell cycle event for a cell (Figure 6.9) treated with $1\times10^{+3}$ nM TPT, starting at $t_{G1} = 280$ (G_1 phase of the newborn cell) min up to $t_3 = 1920$ min. The

dsDNA breaks (induced by TPT) completely disappear at time $t \approx 700$ min, that is, the DNA is successfully repaired. In addition, at time $t \approx 990$ min, the anti-cancer agent TPT (in both forms TPT_L and TPT_H) is totally flushed out of the cell, directly and indirectly via the BCRP/ABCG2 transporter as illustrated in Figure 6.16, which demonstrates the significance of this transporter in drug resistance mechanisms (see Chapters 3 and 4).

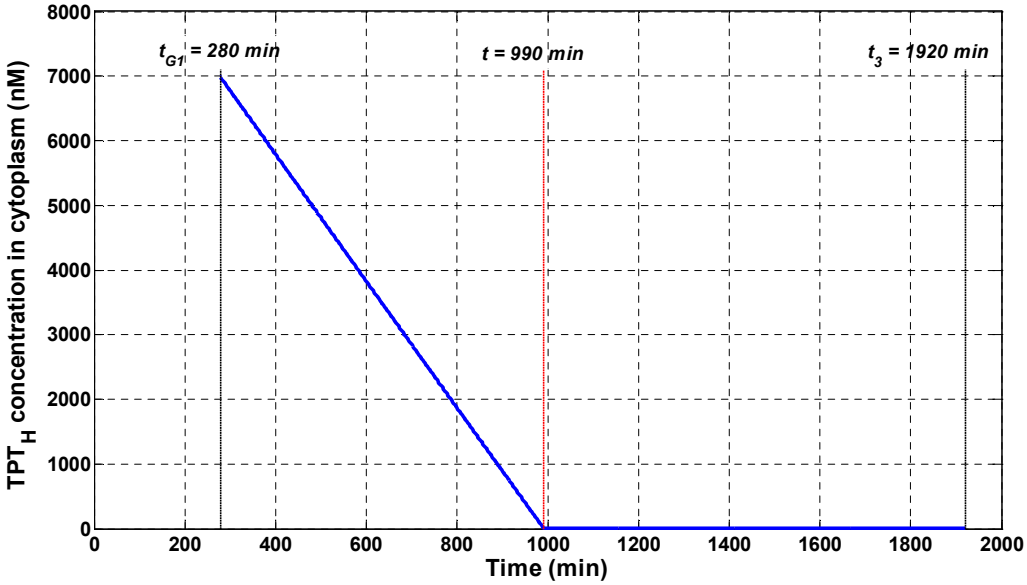


Figure 6.16: Simulated output for TPT_H in the cytoplasm (H_c) from model Equation (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 1\times10^{+3}$ nM TPT during one complete cell cycle.

Figure 6.17 represents the average cytoplasmic Cyclin B1-GFP intensity profile from a G_2 cell responding to a dose $D = 10\times10^{+3}$ nM of TPT. Similar to the previous experiment, tracking of cells begins 20 min after the *washout* event, i.e., at time $t_{start} = 80$ min (see Figure 6.17). In this particular experiment, the first cell division takes place at time $t_B = 620$ min and the complete newborn cell appears at time $t_{G1} = 640$ min. The values for the state variables relating to the cell cycle response component of the coupled model at time t_{G1} are given in Equations (5.16), see Chapter 5. In order to identify the values for the state variables of the full model, the drug kinetics model (including the effect compartment) was simulated to time $t_{G1} = 640$ min (i.e., 580 min following the *washout* event) as shown in Figures 6.18-6.20.

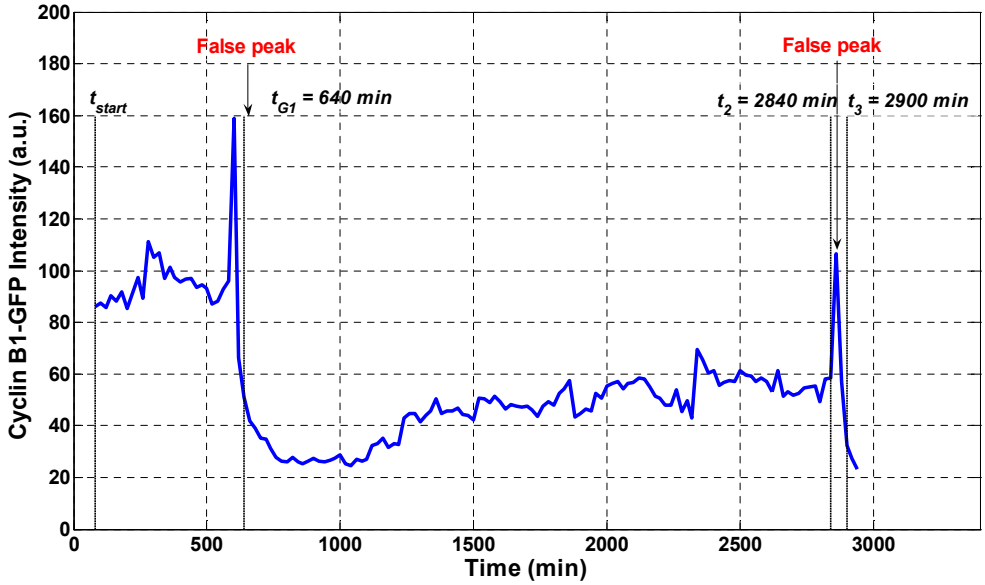


Figure 6.17: Average cytoplasmic Cyclin B1-GFP fluorescence intensity extracted from a typical G_2 cell treated with $10\times10^{+3}$ nM TPT.

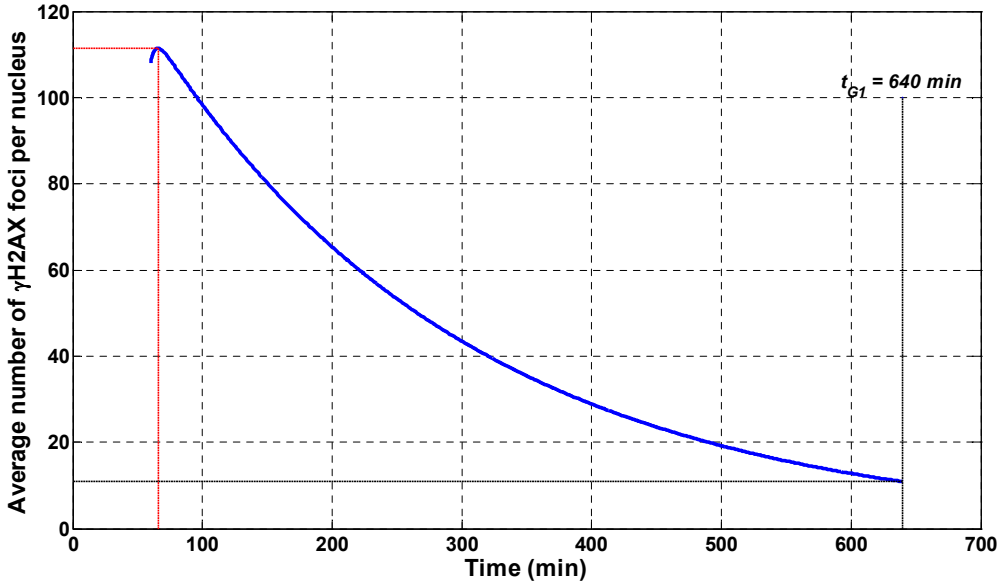


Figure 6.18: Simulated output for the average number of γH2AX foci per nucleus from the model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 10\times10^{+3}$ nM following the *washout* event until time $t = t_{G1}$.

The average number of γH2AX foci per nucleus slightly increases (following the *washout* event) and reaches the maximum number of γH2AX foci per nucleus at around time $t = 65.6667$ min. Then, the number of γH2AX foci starts to exponentially decrease (see Figure 6.18), that is when the cell starts the DNA repair mechanism which does not start immediately after removing the drug from the medium. The concentration of TPT_H in the cytoplasm (H_c) also increases following the *washout* event (Figure 6.19) and reaches a

maximum concentration of $1.3054\times10^{+3}$ nM at time $t = 71.6667$ min as a result of the reversible hydrolysis of L_c to H_c, the irreversible hydrolysis of L_c to H_c (due to the catalytic reaction of the enzyme ALDH) and the dissociation of TPT_L bound to DNA to H_c. Then, H_c starts to decrease and converges to zero (i.e., 0 nM) over time as shown in Figure 6.19. Similar to the previous experiment (i.e., when cells are treated with $D = 10\times10^{+3}$), L_c and L_n start rapidly decreasing after the *washout* event (Figure 6.20). Therefore, the drug (TPT_L and TPT_H) in the cellular compartments (nucleus and cytoplasm) flows to the extracellular region by simple diffusion (only TPT_L) or by the active pumping mechanism (only TPT_H through the BCRP/ABCG2 transporter).

The values for the state variables of the full drug kinetics/cell cycle response model (dynamics) at time $t_{G1} = 640$ min when cells are treated with $10\times10^{+3}$ nM TPT are given by:

$$\begin{gathered}
[\text{CycB}](640) = 0.141396C_2, [\text{CDC2CycB}](640) = 0.992654C_2, \\
[\text{CDC2CycB}_{\text{YP}}](640) = 1.4763C_2, [\text{CDC2CycB}_{\text{YPTP}}](640) = 11.41023C_2, \\
[\text{CDC2CycB}_{\text{TP}}](640) = 1.64121C_2, [\text{CDC2}](640) = 84.479606C_2, \\
[\text{APC}_{\text{active}}](640) = 3.8539C_2, [\text{Wee1}_{\text{active}}](640) = 15C_2, \\
[\text{CDC25}_{\text{active}}](640) = 1.54009C_2, [\text{p21}_{\text{active}}](640) = 3.8539C_2, \\
L_m(640) = 0.4073,\ H_m(640) = 0.2320,\ L_e(640) = 0.3659, \\
H_e(640) = 5.2507,\ L_c(640) = 0,\ H_c(640) = 7.5912\times10^{+3}, \\
L_n(640) = 0 \text{ and } \gamma H2AX(640) = 10.87
\end{gathered} \tag{6.46}$$

where, the state variable γH2AX is dimensionless and the rest of the state variables are measured in nM.

In the Cyclin B1-GFP intensity profile shown in Figure 6.17, the false peak is located at time $t = 2860$ min and the second cell division occurs at time $t = 2880$ min, therefore, a complete newborn cell appears at time $t_3 = 2900$ min. The parameter estimation for the cell treated with $10\times10^{+3}$ nM TPT was performed starting from time $t_{G1} = 640$ min until time $t_3 = 2900$ min excluding the experimental data point which includes the false peak at time $t = 2860$ min. Using Equation (6.45), the scaling factor $C_1 = 0.0613$.

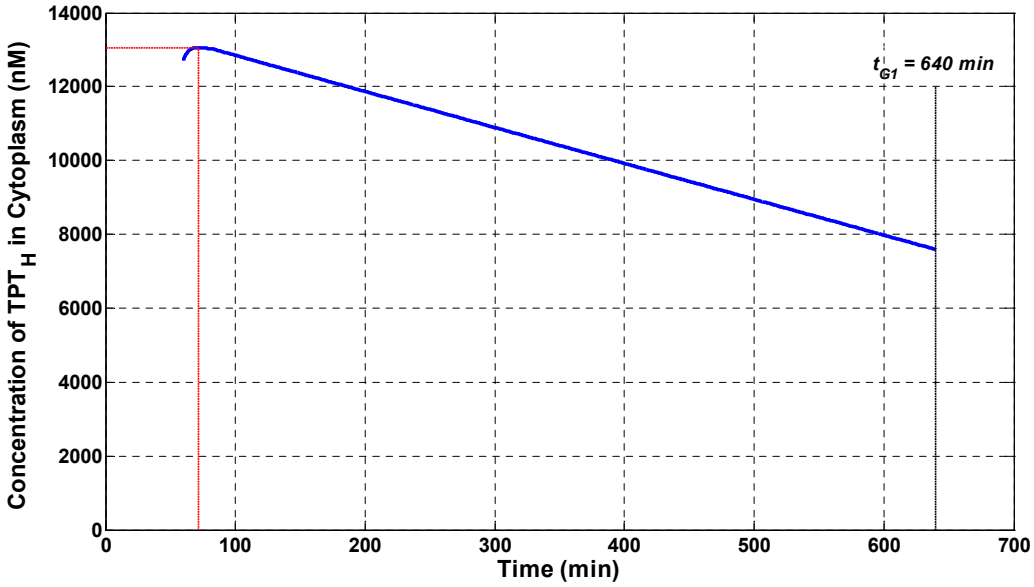


Figure 6.19: Simulated output for TPT_H in the cytoplasm (H_c) from model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 10\times10^{+3}$ nM TPT following the *washout* event until time $t = t_{G1}$.

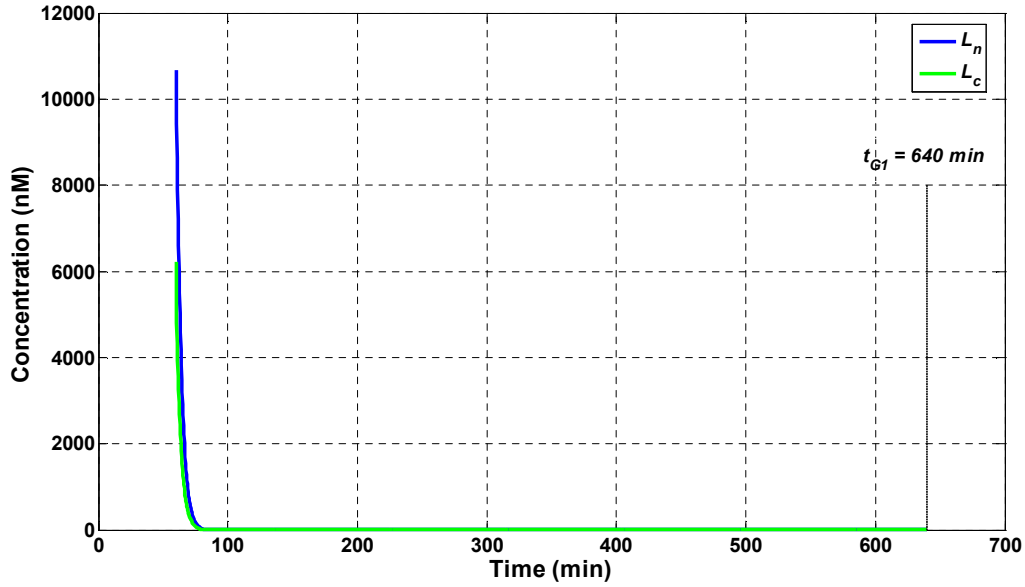


Figure 6.20: Simulated output for L_n (blue) and L_c (green) from model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 10\times10^{+3}$ nM TPT following the *washout* event until time $t = t_{G1}$.

The fitted (estimated) parameter values, with estimates for their confidence levels are presented in Table 6.5, where it is seen that all parameters are well-determined by the Cyclin B1-GFP data. In addition, all of the estimated parameters have low SDLN values corresponding to high confidence in their values. The one exception is the estimated parameter value for $p21_{total}$ due to the low signal-to-noise ratio in the Cyclin B1-GFP data.

The value of $CORI_{mean} = 0.67176$ indicates that the residuals are acceptable and evenly distributed above and below the simulated output of the model and therefore the resulting fit is good.

Table 6.5: Best parameter estimates for the coupled drug kinetics/cell cycle response model Equations (6.1)-(6.14) and state values at $t = t_{G1}$ in Equation (6.46), estimated using live-cell Cyclin B1-GFP data for a cell treated with $10\times10^{+3}$ nM.

Parameter				Value
C_1				0.0613
Parameter	**Value**	**SDLN**	**5%**	**95%**
α (a.u. nM^{-1})	2.21	0.2073	1.57	3.11
C_2	9.23×10^{-1}	0.0237	8.87×10^{-1}	9.59×10^{-1}
β (a.u.)	$3.52\times10^{+1}$	0.0592	$3.19\times10^{+1}$	$3.88\times10^{+1}$
k_{p21off} (min^{-1})	3.08×10^{-2}	0.296	1.89×10^{-2}	5.01×10^{-2}
$p21_{total}$ (nM)	$1.08\times10^{+2}$	0.9677	$2.21\times10^{+1}$	$5.33\times10^{+2}$
RSS = $5.59\times10^{+1}$				

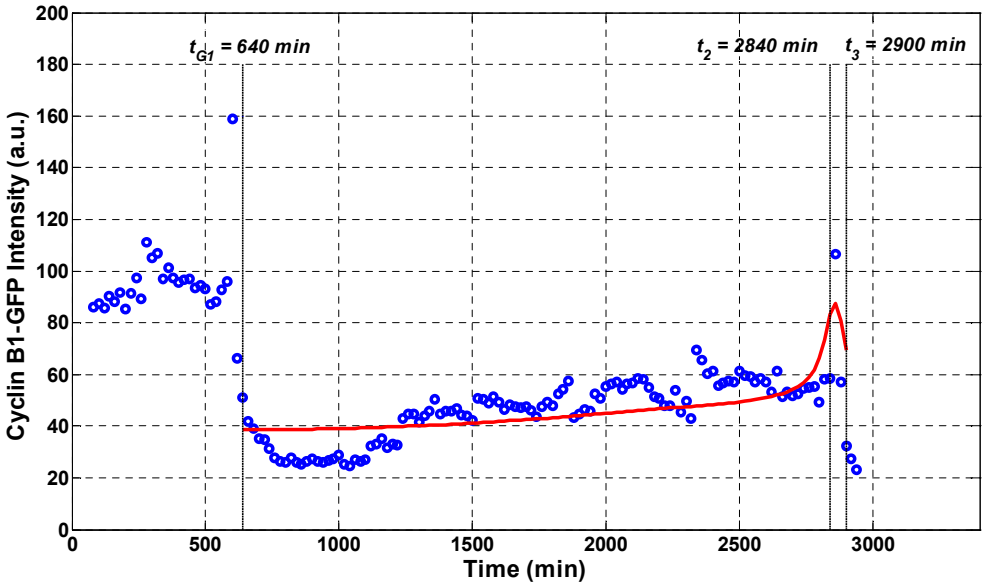


Figure 6.21: Model fit showing model prediction (solid red line) and Cyclin B1-GFP data (blue circles) in the presence of $10\times10^{+3}$ nM TPT. An artefact of morphological changes during the final stages of mitosis is the over-expression of Cyclin B1 fluorescence giving rise to false peaks in the experimental data.

The plots shown in Figure 6.21 demonstrate a typical trace for the growth of a cell treated with $10\times10^{+3}$ nM TPT. The Cyclin B1-GFP data are denoted by the blue circles and the simulated data by the solid red curve. The model fit in Figure 6.21 shows that there is a close reproduction of Cyclin B1-GFP data by simulated output from the coupled drug kinetics/dynamics model with parameters taking values from Tables 6.1-6.3. In addition,

the impact of the false peak (see Figure 6.17) in the experimental data is minimised by the model.

Figure 6.22 shows the normal probability plot for the weighted residuals for the fit in Figure 6.21 of the coupled drug kinetics/cell cycle response model to Cyclin B1-GFP data of a cell treated with $10\times10^{+3}$ nM TPT. The residuals are plotted in ascending order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale. The figure is consistent with the assumption of normally distributed errors (with approximately zero mean and standard deviation of 1), since the graph is approximately linear, which is an indicator for the appropriateness of the standard error σ used for each observation y.

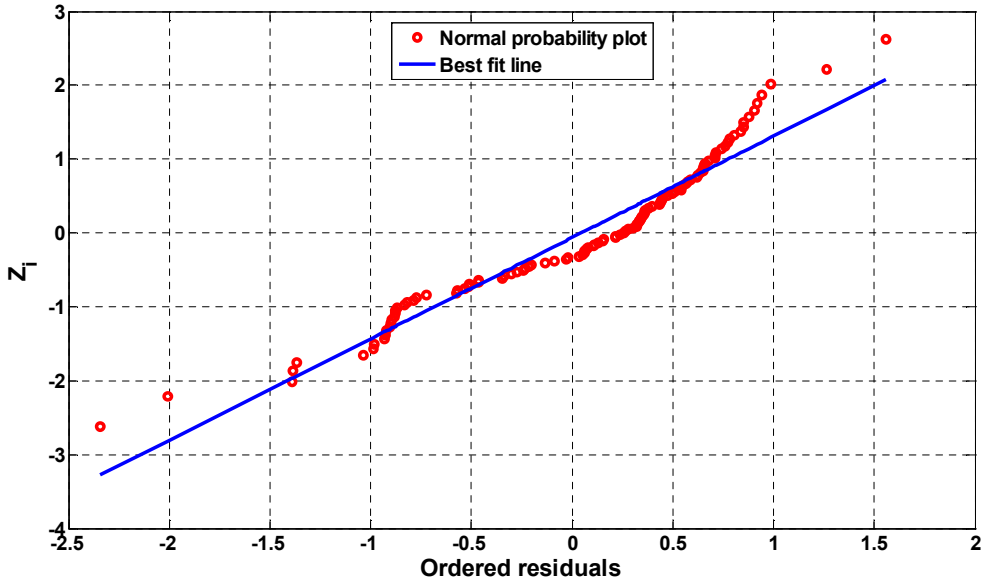


Figure 6.22: A normal probability plot for the residuals in the weighted non-linear least-squares fit of the coupled drug kinetics/cell cycle response model to Cyclin B1-GFP data extracted from typical G_2 cell treated with $10\times10^{+3}$ nM TPT. The residuals are plotted in ascending numerical order on the horizontal axis, against cumulative probabilities in $N(0, 1)$ transformed to a linear scale.

Table 6.6: The estimated correlation matrix for the well-determined parameters for a cell treated with $10\times10^{+3}$ nM TPT.

	α	C_2	β	k_{p21off}	$p21_{total}$
α	1	-0.144	-0.885	0.338	0.811
C_2	-0.144	1	0.202	-0.963	-0.368
β	-0.885	0.202	1	-0.361	-0.694
k_{p21off}	0.338	-0.963	-0.361	1	0.593
$p21_{total}$	0.811	-0.368	-0.694	0.593	1

The estimated correlation matrix for the well-determined parameters is given in Table 6.6. From this table it is seen that the highest correlation is between α and β (-0.855), $p21_{total}$ and α (0.811) and k_{p21off} and C_2 (-0.963). The high correlation of the logarithms of these parameters indicates that it is the product or sum of the parameters that is important for the fitting.

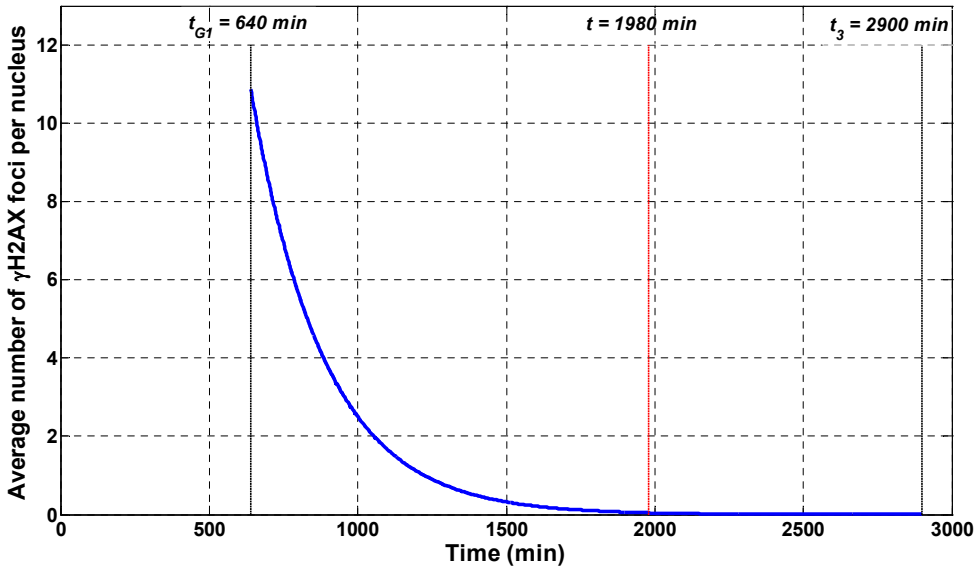


Figure 6.23: Simulated output for the average number of γH2AX foci per nucleus from the model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 10\times10^{+3}$ nM during one complete cell cycle.

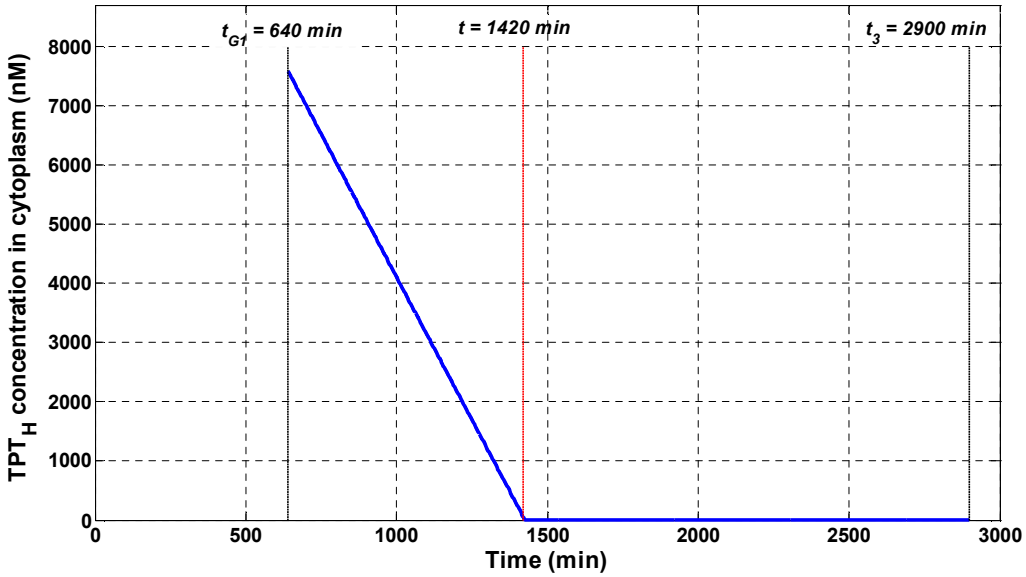


Figure 6.24: Simulated output for TPT_H in the cytoplasm (H_c) from model Equations (6.1), with parameters taking values in Tables 6.1 and 6.2 and the initial dose $D = 10\times10^{+3}$ nM TPT during one complete cell cycle.

Figure 6.23 represents the simulated output for the average number of γH2AX foci per nucleus during the complete cell cycle event of the cell (Figure 6.17) treated with $10\times10^{+3}$ nM TPT, starting at $t_{G1} = 640$ (G_1 phase of the newborn cell) min up to $t_3 = 2900$ min. The dsDNA breaks (induced by TPT) completely disappear at approximately time $t = 1980$ min (i.e., the DNA is successfully repaired, see Figure 6.15). In addition, at approximately time $t = 1420$ min, the drug (in both forms TPT_L and TPT_H) is totally flushed out of the cell, directly (via passive diffusion) and indirectly (via the BCRP/ABCG2 transporter) as illustrated in Figure 6.24, which demonstrates the potential significance of this transporter in drug resistance mechanisms (see Chapters 3 and 4).

In comparison between the results obtained from the parameter estimation for the untreated cell (see Chapter 5), the cell treated with $1\times10^{+3}$ nM TPT (Table 6.3) and the cell treated with $10\times10^{+3}$ nM TPT (Table 6.5), the main difference is the extension in the IMT as a result of the decrease in the *effective* rate at which the CKI $p21^{CIP1/WAF1}$ is switched off, i.e., $C_1k_{p21off}/(1 + k_{dam}\gamma H2AX(t))$. That is, the time required to complete the cell cycle increases when the parameter k_{p21off} decreases due to its inhibitory effect. These results match the expected biological behaviour since both Cyclin B1/CDK1 and the CKI $p21^{CIP1/WAF1}$ are antagonist proteins (see Chapters 3 and 5). Additionally, the estimated parameter values for the constant C_2 (see Section 5.3) for the untreated cell, the cell treated with $1\times10^{+3}$ nM TPT and the cell treated with $10\times10^{+3}$ nM TPT are 9.2038×10^{-1}, 9×10^{-1} and 9.23×10^{-1}, respectively. The estimated values for C_2 (in the three different cells) are relatively close since these cells are from the same cell type (U-2 OS cell line). However, the slight variation among the values for C_2 is caused by the different conditions (i.e., different treatment regimens) these cells are subject to (see Chapter 5). Similar to the parameter C_2, the estimated values for the parameter $p21_{total}$ are of the order of 10^{+2} nM in the untreated cell and the treated cells. This confirms that the concentrations of the cell cycle components are close for cells from the same cell type, however, the variations are caused by the different conditions these cells are subject to. The estimated values for the parameters α and β vary significantly between the untreated cell(s) (see Chapter 5) and the treated cells (Tables 6.3 and 6.5) due to the differences in the experimental set up and the resulting background intensities in the Cyclin B1-GFP data for the control cell (untreated cell) and the treated cells.

6.6 Sensitivity analysis of the coupled drug kinetics/cell cycle response model

It was shown in Chapter 5 that the model variable $CDC2CycB_{TP}$ (the active form of Cyclin B1 protein) exhibits a noticeable variation in simulations when varying the parameter k_{p21off}. In this section, the impact of varying the values for the parameter k_{dam} is considered with respect to the variable $CDC2CycB_{TP}$ in the coupled drug kinetics/dynamics model. The impact of varying the values for the parameter k_{dam} for one complete cycle starting at t_{G1} = 280 min of the cell treated with $1\times10^{+3}$ nM TPT is shown in Figure 6.25. The parameter k_{dam} was varied over the range 1.44×10^{-5} min^{-1}-$1.44\times10^{+1}$ min^{-1}. These values are: (a) 1.44×10^{-5} min^{-1}, (b) 1.44×10^{-2} min^{-1}, (c) 1.44×10^{-1} min^{-1} and , (d) $1.44\times10^{+1}$ min^{-1}.

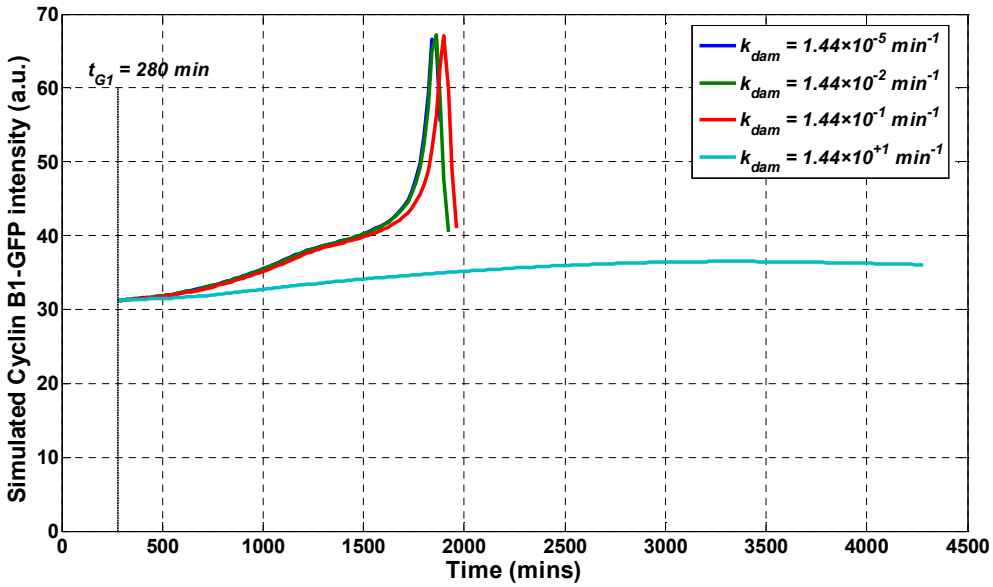


Figure 6.25: Simulated output from the coupled drug kinetics/dynamics model at $t = t_{G1}$ by varying the parameter k_{dam} in the range of 1.44×10^{-5}-$1.44\times10^{+1}$ min^{-1}.

As shown in Figure 6.25, the active form of Cyclin B1 exhibits noticeable variation in simulations when varying the parameter k_{dam}. That is, the activity of Cyclin B1/CDC2 is sensitive to changes in the parameter k_{dam}. Based on the results obtained in Figure 6.25, the time required to complete a cell cycle event increases with an increase in the parameter k_{dam}. This is not surprising since the parameter k_{dam} determines (implicitly) the rate at which the CKI $p21^{CIP1/WAF1}$ is switched off (k_{p21off}). Therefore, an increase in the value of k_{dam} decreases the rate constant k_{p21off} which results in an up-regulation of the CKI

$p21^{CIP1/WAF1}$. Further increase in the value of the parameter k_{dam} can result in a cell cycle arrest (see Chapter 3) as shown in Figure 6.25.

6.7 Summary

Chapter 6 presents a *novel* coupling of the extended drug kinetics model and the extended cell cycle model to describe the response of the growth of single human osteosarcoma cells in the presence of $1\times10^{+3}$ and $10\times10^{+3}$ nM TPT. Modelling the dynamic effect (response) of TPT on human osteosarcoma cells was achieved by extending the drug kinetics model presented in Chapter 4 by introducing an effect compartment to represent γH2AX which is linked to the CKI $p21^{CIP1/WAF1}$ in the extended cell cycle model. The model proposed in this experiment was found to be structurally globally identifiable with respect to the high content screening experiments (DNA damage data) and Cyclin B1-GFP microscopy experiments used to collect data. The corresponding simulations show a good qualitative representation of the experimental data. However, the fits for the cell treated with $1\times10^{+3}$ nM TPT does not show close reproduction of the data by simulated output from the coupled model due to the noise in the data.

7 Conclusions

In this final chapter, the significant outcomes of the work presented in this thesis will be identified and discussed. In all models developed in this thesis, mathematical modelling of biological systems has proven to be useful, particularly when investigating parts of the system that are not directly measured (i.e., that are not physically accessible).

In Chapter 4, a state-space model of the *in vitro* kinetics of the anti-cancer agent TPT has been extended from a previously published model. The postulated model takes into account the effect of the ALDH enzyme and the elimination of the drug (TPT_H) from the cytoplasm via the efflux pump (ABCG2/BCRP drug transporter). The extended drug kinetics model was reduced by assuming that the binding kinetics of both the ALDH enzyme and the BCRP/ABCG2 transporter are saturable and therefore, PSSA applies. Structural identifiability, which is an *a priori* analysis for parameter estimation, was performed using the computer algebra system MATHEMATICA, this tool has proven useful for symbolic computations mainly when they become algebraically complicated. The identifiability analysis demonstrates that all of the unknown model parameters (of the reduced drug kinetics model) are uniquely determined by the output structure corresponding to the real experiment. This ensures that each model parameter can be uniquely determined from perfect, noise-free and continuous observations.

Parameter estimation was performed using the commercial simulation software package FACSIMILE. The optimisation method used in this computer-modelling tool to obtain parameter estimates involves minimisation of the weighted least-square criterion. One advantage of this tool is the robust and reliable numerical integrator that is capable of handling extremely stiff systems efficiently. Initially, three different curve fittings were conducted for average, low loading and high loading cells. Model simulations were compared with live human breast cancer cell (MCF-7 cell line) data and found to give good qualitative agreement. Based on the results obtained from the three curve fittings, the averaged data from 13 individual cells, the high loading cell and the low loading cell, the active pumping mechanism resulting from BCRP/ABCG2 transporter is a key factor in

resisting the anti-cancer agent TPT. To ensure that the results are appropriate and to investigate the heterogeneity in response to the drug, the kinetic model was fitted to a high loading cell and a low loading cell simultaneously by assuming that variations across cells are only due to differences between the parameter relating to the ALDH enzyme and BCRP/ABCG2 transporter. Similar to the individual fittings, simulated output from the drug kinetics model accurately reproduces the behaviour of the underlying biological system being represented. In addition, the results obtained from the model confirm that the efflux pumping mechanism plays a key role in the resistance of anti-cancer drugs. Further analysis showed that model fits were sensitive to the binding affinity of the drug to the transporter (ABCG2/BCRP) and to the binding affinity of the drug to the enzyme ALDH, confirming that both the transporter and the enzyme resist TPT.

One of the motivations for mathematical modelling is the ability to perform *in silico* estimations and predictions about future system behaviour in addition to the relationship between the target binding and the dose. The impact of the dose was measured by calculating the AUC (integral) of the model variable L_n (TPT_L bound to DNA target) corresponding to the initial dose over the first hour. It was found that the high loader and low loader respond differently to the dose and therefore, this variable (i.e., L_n) is an important factor in modelling the effect of the drug (dynamic modelling).

In Chapter 5, a mathematical model describing the growth of single human cells in the absence of TPT was presented. The cell cycle model described in Chapter 5 is a *novel* extension of a previously published model for *Xenopus laevis* embryos to account for the inhibition effect of the CKI $p21^{CIP1/WAF1}$ (under the transcriptional control of p53). For a realistic cell cycle model, the effect of the CKI $p21^{CIP1/WAF1}$ must be taken into account since in eukaryotic cells, this inhibitor plays an important role in stabilising the G_1 state. A unique feature of the extended cell cycle model is that it takes into consideration the variations (in the interaction rate constants and concentrations of the cell cycle components) between different cell types of the same organism and within the same cell type under different conditions. This feature was added by introducing the scaling factors C_1 and C_2.

The structural identifiability analysis of the extended cell cycle model demonstrates that all of the unknown model parameters are uniquely determined by the output structure corresponding to the real experiment. Similar to identifiability analysis of the drug kinetics model, the structural identifiability analysis of the extended cell cycle model was performed using MATHEMATICA. The unknown model parameters were estimated (using FACSIMILE) from live human osteosarcoma cells (Cyclin B1-GFP data) of a control (untreated cell) for one complete cell cycle event of a newborn cell starting at G_1 of the cycle, to investigate the effects of the CKI $p21^{CIP1/WAF1}$ on the progression of the cell cycle traverse. The model fit (of the untreated cell) shows that there is close reproduction of the experimental data by simulated output from the extended cell cycle model. In addition, the false peak (spike) in the experimental data (due to rounding up of cell in the final stages of the cell cycle) is minimised by the extended model. Sensitivity analysis of the extended cell cycle model showed that the model variable $CDC2CycB_{TP}$ (the active form of Cyclin B1) is sensitive to small variations in the values for parameters relating to the CKI $p21^{CIP1/WAF1}$. That is, the increase in the activity of the CKI results in an extension in the IMT and possible cell cycle arrest. Therefore, the CKI $p21^{CIP1/WAF1}$ is an important protein in regulating the cell cycle traverse.

Chapter 6 presents a *novel* coupling of the extended drug kinetics model (in Chapter 4) and the extended cell cycle model (in Chapter 5) to describe the response of the growth of single human osteosarcoma cells (U-2 OS cells) in the presence of TPT. Coupling the models in this way, rather than using separate approaches, allows examination of how the drug perturbs the cell cycle. Modelling the dynamic effect (response) of TPT on human osteosarcoma cells was achieved by extending the drug kinetics model presented in Chapter 4 by introducing an effect compartment to represent the DNA damage (γH2AX) which is linked to the CKI $p21^{CIP1/WAF1}$ in the extended cell cycle model (presented in Chapter 5).

The structural identifiability analysis for the coupled drug kinetics/dynamics model was performed using MATHEMATICA. The state-space model proposed in this experiment was found to be structurally globally identifiable with respect to the high content screening experiments (DNA damage data) and Cyclin B1-GFP microscopy experiments used to collect data. The predicted values (using FACSIMILE) for the number of γH2AX foci per

nucleus are identical to the experimental values. The results obtained from this chapter also confirm drug resistance to TPT resulting from the BCRP/ABCG2 transporter and the enzyme ALDH following the *washout* event. In addition, it has predicted that extension in the IMT as a result of the DNA damage is due to the up-regulation of the CKI $p21^{CIP1/WAF1}$. It was illustrated that this coupling has proven to be useful in studying how the drug perturbs the cell cycle. In addition, a robust and validated version of the current model could be used to design drug treatment regimens, i.e., treatment plans that specify dosage, schedule and duration of treatment.

7.1 Future work

The present drug kinetics model does not address inter-cellular heterogeneity across the entire population of cells (i.e., the 13 individual cells) but the model was used to compare two different responses to the drug. Future work can include an extension of the single-cell model (presented in Chapter 4) to a multi-cell model to account for inter-cellular heterogeneity which is similar to the approach used by Cheung *et al.* [48]. In recent years, a wide assortment of BCRP inhibitors (chemosensitisers) have been identified, which can also be considered in future research on the extended drug kinetics model.

The current extended cell cycle model includes the basic feature of the cell cycle components. However, other components such as p53 tumour suppressor gene can be incorporated into the model. One current limitation is the available microscopy data, and suggestions in this respect include acquiring further data from the cell in addition to the activity of Cyclin B1 protein.

The cell cycle dynamics data provided by the biological collaborators (at Cardiff) are for live human osteosarcoma cells. The extended drug kinetics model has been validated to a certain degree from live human breast cancer cells. For a more robust drug kinetics/cell cycle response model, it is more appropriate to have coupling from the same cell type. Future work includes extending the current *in vitro* drug kinetics/dynamics model for *in vivo* applications using physiologically-based pharmacokinetic/pharmacodynamic

(PBPK/PD) modelling techniques. PBPK/PD modelling approaches are very powerful tools which account for the anatomical and physiological structure of the body. In addition, they assist in determining relationships between drug levels, dosing regimens and response to drug *in vivo* which are critical to the success of clinical trials.

References

[1] K. Tsaioun, S.A. Kates, (Eds.), ADMET for Medicinal Chemists: A Practical Guide, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2011.

[2] M. Abou-Gharbia, Discovery of innovative small molecule therapeutics, Journal of Medicinal Chemistry 52 (2009) 2-9.

[3] M. Dickson, J.P. Gagnon, Key factors in the rising cost of new drug discovery and development, Nature Reviews Drug Discovery 3 (2004) 417-429.

[4] E.F. Schmid, D.A. Smith, Pharmaceutical R&D in the spotlight: why is there still unmet medical need?, Drug Discov Today 12 (2007) 998-1006.

[5] D.A. Smith, E.F. Schmid, Drug withdrawals and the lessons within, Current Opinion in Drug Discovery & Development 9 (2006) 38-46.

[6] E.F. Schmid, D.A. Smith, Keynote review: Is declining innovation in the pharmaceutical industry a myth?, Drug Discov Today 10 (2005) 1031-1039.

[7] A.E. Nassar, A.M. Kamel, C. Clarimont, Improving the decision-making process in structural modification of drug candidates: reducing toxicity, Drug Discovery Today 9 (2004) 1055-1064.

[8] C. Csajka, D. Verotta, Pharmacokinetic-pharmacodynamic modelling: History and perspectives, Journal of Pharmacokinetics and Pharmacodynamics 33 (2006) 227-279.

[9] L. Aarons, M.O. Karlsson, F. Mentré, F. Rombout, J.-L. Steimer, A. van Peer, COST B15 Experts, Role of modelling and simulation in Phase I drug development, European Journal of Pharmaceutical Sciences 13 (2001) 115-122.

[10] L.B. Sheiner, J.L. Steimer, Pharmacokinetic/pharmacodynamic modeling in drug development, Annual Review of Pharmacology and Toxicology 40 (2000) 67-95.

[11] J.Y. Chien, S. Friedrich, M.A. Heathman, D.P. de Alwis, V. Sinha, Pharmacokinetics/pharmacodynamics and the stages of drug development: role of modeling and simulation, The AAPS Journal 7 (2005) E544-E559.

[12] R.J. Errington, S.M. Ameer-beg, B. Vojnovic, L.H. Patterson, M. Zloh, P.J. Smith, Advanced microscopy solutions for monitoring the kinetics and dynamics of drug-DNA targeting in living cells, Advanced Drug Delivery Reviews 57 (2005) 153-167.

[13] N.D. Evans, R.J. Errington, M.J. Chapman, P.J. Smith, M.J. Chappell, K.R. Godfrey, Compartmental modelling of the uptake kinetics of the anti-cancer agent topotecan in human breast cancer cells, International Journal of Adaptive Control and Signal Processing 19 (2005) 395-417.

[14] J.R. Pomerening, E.D. Sontag, J.E. Ferrell Jr., Building a cell cycle oscillator: hysteresis and bistability in the activation of Cdc2, Nature Cell Biology 5 (2003) 346-351.

[15] N.D. Evans, R.J. Errington, M. Shelley, G.P. Feeney, M.J. Chapman, K.R. Godfrey, P.J. Smith, M.J. Chappell, A mathematical model for the in vitro kinetics of the anti-cancer agent topotecan, Mathematical Biosciences 189 (2004) 185-217.

[16] J.D. Murray, Mathematical Biology I: An Introduction, 3rd Ed., Springer, New York, USA, 2002.

[17] L. Alberghina, H.V. Westerhoff, (Eds.), Systems Biology: Definitions and Perspectives (Topics in Current Genetics), Springer, 2005.

[18] C.M. Close, D.K. Frederick, Modeling and Analysis of Dynamic Systems, 2nd Ed., Houghton Mifflin Company, Boston, MA, USA, 1993.

[19] E.A. Bender, An Introduction to Mathematical Modeling, John Wiley & Sons, Inc., New York, NY, USA, 1978.

[20] V.Z. Marmarelis, Nonlinear Dynamic Modeling of Physiological Systems, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2004.

[21] K.J. Himmelstein, R.J. Lutz, A review of the applications of physiologically based pharmacokinetic modeling, Journal of Pharmacokinetics and Biopharmaceutics 7 (1979) 127-145.

[22] D. Voet, J.G. Voet, Biochemistry, 2nd Ed., John Wiley & Sons, Inc., New York, USA, 1995.

[23] J.A. Jacquez, Compartmental Analysis in Biology and Medicine, 3rd Ed., BioMedware, Ann Arbor, MI, USA, 1996.

[24] K.R. Godfrey, Compartmental Models and their Application, Academic Press, London, UK, 1983.

[25] K.R. Godfrey, J.J. DiStefano III, Identifiability of model parameters, in: E. Walter (Ed.), Identifiability of Parametric Models, Pergamon Press, Oxford, 1987, p.1 (Chapter 1).

[26] J.J. DiStefano III, The modeling methodology forum: an expanded department, American Journal of Physiology - Heart and Circulatory Physiology 248 (1985) H155-H156.

[27] F.G. Heineken, H.M. Tsuchiya, R. Aris, On the mathematical status of the pseudo-steady state hypothesis of biochemical kinetics, Mathematical Biosciences 1 (1967) 95-113.

[28] A.L. Lehninger, D.L. Nelson, M.M. Cox, Lehninger Principles of Biochemistry, 4th Ed., W.H. Freeman, New York, USA, 2005.

[29] J.G. Wagner, Kinetics of pharmacologic response I. Proposed relationships between response and drug concentration in the intact animal and man, Journal of Theoretical Biology 20 (1968) 173-201.

[30] L.B. Sheiner, D.R. Stanski, S. Vozeh, R.D. Miller, J. Ham, Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine, Clinical Pharmacology and Therapeutics 25 (1979) 358-371.

[31] C.J. Hull, H.B.H. Van Beem, K. McLeod, A. Sibbald, M.J. Watson, A pharmacodynamic model for pancuronium, British Journal of Anaesthesia 50 (1978) 1113-1123.

[32] B. Whiting, A.W. Kelman, Modeling of drug response in individual subjects, Journal of Pharmacokinetics and Biopharmaceutics 8 (1980) 115-130.

[33] R. Bellman, K.J. Åström, On structural identifiability, Mathematical Biosciences 7 (1970) 329-339.

[34] E. Walter, (Ed.), Identifiability of Parametric Models, Pergamon Press, Oxford, 1987.

[35] N.D. Evans, M.J. Chapman, M.J. Chappell, K.R. Godfrey, Identifiability of uncontrolled nonlinear rational systems, Automatica 38 (2002) 1799-1805.

[36] E. Walter, Y. Lecourtier, Unidentifiable compartmental models: what to do?, Mathematical Biosciences 56 (1981) 1-25.

[37] S. Vajda, K.R. Godfrey, H. Rabitz, Similarity transformation approach to identifiability analysis of nonlinear compartmental models, Mathematical Biosciences 93 (1989) 217-248.

[38] E.T. Tunali, T.-J. Tarn, New results for identifiability of nonlinear systems, IEEE Transactions on Automatic Control 32 (1987) 146-154.

[39] M. Fliess, S.T. Glad, An algebraic approach to linear and nonlinear control, in: H.L. Trentelman and J. C. Willems (Eds.), Essays on Control: Perspectives in the Theory

and its Applications, Birkhäuser, Boston, 1987, Vol. 14 of Progress in Systems and Control Theory.

[40] L. Ljung, T. Glad, On global identifiability for arbitrary model parameterizations, Automatica 30 (1994) 265-276.

[41] H. Pohjanpalo, System identifiability based on the power series expansion of the solution, Mathematical Biosciences 41 (1978) 21-33.

[42] M.J. Chappell, K.R. Godfrey, S. Vajda, Global identifiability of the parameters of nonlinear systems with specified inputs: a comparison of methods, Mathematical Biosciences 102 (1990) 41-73.

[43] J.A. Jacquez, Identifiability: the first step in parameter estimation, Federation Proceedings 46 (1987) 2477-2480.

[44] J.A. Jacquez, P. Greif, Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design, Mathematical Biosciences 77 (1985) 201-227.

[45] K.R. Godfrey, W.R. Fitch, The deterministic identifiability of nonlinear pharmacokinetic models, Journal of Pharmacokinetics and Biopharmaceutics 12 (1984) 177-191.

[46] S. Vajda, Structural identifiability of linear, bilinear, polynomial and rational systems, 9th IFAC World Congress, Budapest, Hungary, 1982.

[47] S. Vajda, Identifiability of polynomial systems: structural and numerical aspects, in: E. Walter (Ed.), Identifiability of Parametric Models, Pergamon Press, Oxford, 1987, p. 42 (Chapter 4).

[48] S.Y.A. Cheung, N.D. Evans, M.J. Chappell, K.R. Godfrey, P.J. Smith, R.J. Errington, Exploration of the intercellular heterogeneity of topotecan uptake into human breast cancer cells through compartmental modelling, Mathematical Biosciences 213 (2008) 119-134.

[49] AEA Technology Harwell Laboratory, Facsimile (Version 4.0) Technical Reference Manual, Didcot, Oxfordshire, UK, 1995.

[50] AEA Technology Harwell Laboratory, Facsimile (Version 4.0) User Guide, Didcot, Oxfordshire, UK, 1995.

[51] Y. Bard, Nonlinear Parameter Estimation, Academic Press, NY, USA, 1974.

[52] D.C. Montgomery, G.C. Runger, Applied Statistics and Probability for Engineers, 4th Ed., John Wiley & Sons, Inc., Hoboken, NJ, USA, 2007.

[53] J. O'Leary, F.M. Muggia, Camptothecins: a review of their development and schedules of administration, The European Journal of Cancer 34 (1998) 1500-1508.

[54] C. Bailly, Topoisomerase I poisons and suppressors as anticancer drugs, Current Medicinal Chemistry 7 (2000) 39-58.

[55] M.E. Wall, M.C. Wani, Camptothecin and taxol: discovery to clinic--thirteenth Bruce F. Cain Memorial Award Lecture, Cancer Research 55 (1995) 753-760.

[56] M.E. Wall, M.C. Wani, C.E. Cook, K.H. Palmer, A.T. McPhail, G.A. Sim, Plant Antitumor Agents I. The isolation and structure of camptothecin, a novel alkaloidal leukemia and tumor inhibitor from camptotheca acuminata, Journal of the American Chemical Society 88 (1966) 3888-3890.

[57] A.M. Abang, The clinical pharmacology of topoisomerase I inhibitors, Seminars in Hematology 35 (1998) 13-21.

[58] W.E. Ross, DNA topoisomerases as targets for cancer therapy, Biochemical Pharmacology 34 (1985) 4191-4195.

[59] Y.-H. Hsiang, L.F. Liu, Identification of mammalian DNA topoisomerase I as an intracellular target of the anticancer drug camptothecin, Cancer Research 48 (1988) 1722-1726.

[60] Y.-H. Hsiang, R. Hertzberg, S. Hecht, L.F. Liu, Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I, The Journal of Biological Chemistry 260 (1985) 14873-14878.

[61] T. Andoh, K. Ishii, Y. Suzuki, Y. Ikegami, Y. Kusunoki, Y. Takemoto, K. Okada, Characterization of a mammalian mutant with a camptothecin-resistant DNA topoisomerase I, Proceedings of the National Academy of Sciences of the United States of America 84 (1987) 5565-5569.

[62] E. Kjeldsen, B.J. Bonven, T. Andoh, K. Ishii, K. Okada, L. Bolund, O. Westergaard, Characterization of a camptothecin-resistant human DNA topoisomerase I, Journal of Biological Chemistry 263 (1988) 3912-3916.

[63] M.J. Luzzio, J.M. Besterman, D.L. Emerson, M.G. Evans, K. Lackey, P.L. Leitner, G. McIntyre, B. Morton, P.L. Myers, M. Peel, J.M. Sisco, D.D. Sternbach, W.Q. Tong, A. Truesdale, D.E. Uehling, A. Vuong, J. Yates, Synthesis and antitumor activity of novel water soluble derivatives of camptothecin as specific inhibitors of topoisomerase I, Journal of Medicinal Chemistry 38 (1995) 395-401.

[64] X. Wang, L.K. Wang, W.D. Kingsbury, R.K. Johnson, S.M. Hecht, Differential effects of camptothecin derivatives on topoisomerase I-mediated DNA structure modification, Biochemistry 37 (1998) 9399-9408.

[65] M. Sugimori, A. Ejima, S. Ohsuki, K. Uoto, I. Mitsui, Y. Kawato, Y. Hirota, K. Sato, H. Terasawa, Synthesis and antitumor activity of ring A- and F-modified hexacyclic camptothecin analogues, Journal of Medicinal Chemistry 41 (1998) 2308-2318.

[66] O. Lavergne, L. Lesueur-Ginot, F. Pla Rodas, P.G. Kasprzyk, J. Pommier, D. Demarquay, G. Prévost, G. Ulibarri, A. Rolland, A.-M. Schiano-Liberatore, J. Harnett, D. Pons, J. Camara, D.C.H. Bigg, Homocamptothecins: synthesis and antitumor activity of novel E-ring-modified camptothecin analogues, Journal of Medicinal Chemistry 41 (1998) 5410-5419.

[67] R. Garcia-Carbonero, J.G. Supko, Current perspectives on the clinical experience, pharmacology, and continued development of the camptothecins, Clinical Cancer Research 8 (2002) 641-661.

[68] V.M.M. Herben, W.W. ten Bokkel Huinink, J.H.M. Schellens, J.H. Beijnen, Clinical pharmacokinetics of camptothecin topoisomerase I inhibitors, Pharmacy World & Science 20 (1998) 161-172.

[69] C.H. Takimoto, J. Wright, S.G. Arbuck, Clinical applications of the camptothecins, Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1400 (1998) 107-119.

[70] L. Iyer, M.J. Ratain, Clinical pharmacology of camptothecins, Cancer Chemotherapy and Pharmacology 42 (1998) S31-S43.

[71] N. Osheroff, Biochemical basis for the interactions of type I and type II topoisomerases with DNA, Pharmacology & Therapeutics 41 (1989) 223-241.

[72] M. Gellert, DNA topoisomerases, Annual Review of Biochemistry 50 (1981) 879-910.

[73] J.C. Wang, DNA topoisomerases, Annual Review of Biochemistry 54 (1985) 665-697.

[74] H.P. Vosberg, DNA topoisomerases: enzymes that control DNA conformation, Current Topics in Microbiology and Immunology 114 (1985) 19-102.

[75] T. Uemura, M. Yanagida, Isolation of type I and II DNA topoisomerase mutants from fission yeast: single and double mutants show different phenotypes in cell growth and chromatin organization, The EMBO Journal 3 (1984) 1737-1744.

[76] T. Goto, J.C. Wang, Cloning of yeast TOP1, the gene encoding DNA topoisomerase I, and construction of mutants defective in both DNA topoisomerase I and DNA topoisomerase II, Proceedings of the National Academy of Sciences of the United States of America 82 (1985) 7178-7182.

[77] R.M. Snapka, Topoisomerase inhibitors can selectively interfere with different stages of simian virus 40 DNA replication, Molecular and Cellular Biology 6 (1986) 4221-4227.

[78] S.J. Brill, S. DiNardo, K. Voelkel-Meiman, R. Sternglanz, Need for DNA topoisomerase activity as a swivel for DNA replication for transcription of ribosomal RNA, Nature 326 (1987) 414-416.

[79] L. Yang, M.S. Wold, J.J. Li, T.J. Kelly, L.F. Liu, Roles of DNA topoisomerases in simian virus 40 DNA replication in vitro, Proceedings of the National Academy of Sciences of the United States of America 84 (1987) 950-954.

[80] B.D. Halligan, J.L. Davis, K.A. Edwards, L.F. Liu, Intra- and intermolecular strand transfer by HeLa DNA topoisomerase I, Journal of Biological Chemistry 257 (1982) 3995-4000.

[81] P. Bullock, J.J. Champoux, M. Botchan, Association of crossover points with topoisomerase I cleavage sites: a model for nonhomologous recombination, Science 230 (1985) 954-958.

[82] W.K. McCoubrey, J.J. Champoux, The role of single-strand breaks in the catenation reaction catalyzed by the rat type I topoisomerase, Journal of Biological Chemistry 261 (1986) 5130-5137.

[83] G.G. Maul, B.T. French, W.J. van Venrooij, S.A. Jimenez, Topoisomerase I identified by scleroderma 70 antisera: enrichment of topoisomerase I at the centromere in mouse mitotic cells before anaphase, Proceedings of the National Academy of Sciences of the United States of America 83 (1986) 5145-5149.

[84] G. Fleischmann, G. Pflugfelder, E.K. Steiner, K. Javaherian, G.C. Howard, J.C. Wang, S.C. Elgin, Drosophila DNA topoisomerase I is associated with transcriptionally active regions of the genome, Proceedings of the National Academy of Sciences of the United States of America 81 (1984) 6958-6962.

[85] M.T. Muller, W.P. Pfund, V.B. Mehta, D.K. Trask, Eukaryotic type I topoisomerase is enriched in the nucleolus and catalytically active on ribosomal DNA, The EMBO Journal 4 (1985) 1237-1243.

[86] B.J. Bonven, E. Gocke, O. Westergaard, A high affinity topoisomerase I binding sequence is clustered at DNAase I hypersensitive sites in Tetrahymena R-chromatin, Cell 41 (1985) 541-551.

[87] D.S. Gilmour, G. Pflugfelder, J.C. Wang, J.T. Lis, Topoisomerase I interacts with transcribed regions in Drosophila cells, Cell 44 (1986) 401-407.

[88] L.C. Garg, S. DiAngelo, S.T. Jacob, Role of DNA topoisomerase I in the transcription of supercoiled rRNA gene, Proceedings of the National Academy of Sciences of the United States of America 84 (1987) 3185-3188.

[89] L.F. Liu, J.C. Wang, Supercoiling of the DNA template during transcription, Proceedings of the National Academy of Sciences of the United States of America 84 (1987) 7024-7027.

[90] H. Hashimoto, S. Chatterjee, N.A. Berger, Mutagenic activity of topoisomerase I inhibitors, Clinical Cancer Research 1 (1995) 369-376.

[91] J.J. Champoux, DNA is linked to the rat liver DNA nicking-closing enzyme by a phosphodiester bond to tyrosine, Journal of Biological Chemistry 256 (1981) 4805-4809.

[92] L. Stewart, M.R. Redinbo, X. Qiu, W.G.J. Hol, J.J. Champoux, A model for the mechanism of human topoisomerase I, Science 279 (1998) 1534-1541.

[93] M.E. Wall, M.C. Wani, Antineoplastic agents from plants, Annual Review of Pharmacology and Toxicology 17 (1977) 117-132.

[94] M.C. Wani, P.E. Ronman, J.T. Lindley, M.E. Wall, Plant antitumor agents. 18. Synthesis and biological activity of camptothecin analogues, Journal of Medicinal Chemistry 23 (1980) 554-560.

[95] R.P. Hertzberg, M.J. Caranfa, K.G. Holden, D.R. Jakas, G. Gallagher, M.R. Mattern, S.-M. Mong, J.O. Bartus, R.K. Johnson, W.D. Kingsbury, Modification of the hydroxy lactone ring of camptothecin: inhibition of mammalian topoisomerase I and biological activity, Journal of Medicinal Chemistry 32 (1989) 715-720.

[96] Y.-H. Hsiang, L.F. Liu, M.E. Wall, M.C. Wani, A.W. Nicholas, G. Manikumar, S. Kirschenbaum, R. Silber, M. Potmesil, DNA topoisomerase I-mediated DNA cleavage and cytotoxicity of camptothecin analogues, Cancer Research 49 (1989) 4385-4389.

[97] C. Jaxel, K.W. Kohn, M.C. Wani, M.E. Wall, Y. Pommier, Structure-activity study of the actions of camptothecin derivatives on mammalian topoisomerase I: evidence

for a specific receptor site and a relation to antitumor activity, Cancer Research 49 (1989) 1465-1469.

[98] W.J.M. Underberg, R.M.J. Goossen, B.R. Smith, J.H. Beijnen, Equilibrium kinetics of the new experimental anti-tumour compound SK&F 104864-A in aqueous solution, Journal of Pharmaceutical and Biomedical Analysis 8 (1990) 681-683.

[99] J. Fassberg, V.J. Stella, A kinetic and mechanistic study of the hydrolysis of camptothecin and some analogues, Journal of Pharmaceutical Sciences 81 (1992) 676-684.

[100] I. Gryczynski, Z. Gryczynski, J.R. Lakowicz, D.Z. Yang, T.G. Burke, Fluorescence spectral properties of the anticancer drug topotecan by steady-state and frequency domain fluorometry with one-photon and multi-photon excitation, Photochemistry and Photobiology 69 (1999) 421-428.

[101] S. Yao, D. Murali, P. Seetharamulu, K. Haridas, P.N.V. Petluru, D.G. Reddy, F.H. Hausheer, Topotecan lactone selectively binds to double- and single-stranded DNA in the absence of topoisomerase I, Cancer Research 58 (1998) 3782-3786.

[102] T.G. Burke, C.B. Munshi, Z.H. Mi, Y. Jiang, The important role of albumin in determining the relative human blood stabilities of the camptothecin anticancer drugs, Journal of Pharmaceutical Sciences 84 (1995) 518-519.

[103] H. Rosing, D.M. van Zomeren, E. Doyle, W.W. ten Bokkel Huinink, J.H.M. Schellens, A. Bult, J.H. Beijnen, Quantification of topotecan and its metabolite N-desmethyltopotecan in human plasma, urine and faeces by high-performance liquid chromatographic methods, Journal of Chromatography B: Biomedical Sciences and Applications 727 (1999) 191-203.

[104] T.G. Burke, Chemistry of the camptothecins in the bloodstream. Drug stabilization and optimization of activity, Annals of the New York Academy of Sciences 803 (1996) 29-31.

[105] R.P. Hertzberg, M.J. Caranfa, K.G. Holden, D.R. Jakas, G. Gallagher, M.R. Mattern, S.M. Mong, J.O. Bartus, R.K. Johnson, W.D. Kingsbury, Modification of the Hydroxy Lactone Ring of Camptothecin - Inhibition of Mammalian Topoisomerase-I and Biological-Activity, Journal of Medicinal Chemistry 32 (1989) 715-720.

[106] S.E. Porter, J.J. Champoux, The basis for camptothecin enhancement of DNA breakage by eukaryotic topoisomerase I, Nucleic Acids Research 17 (1989) 8521-8532.

[107] M.J. Chappell, N.D. Evans, R.J. Errington, I.A. Khan, L. Campbell, R. Ali, K.R. Godfrey, P.J. Smith, A coupled drug kinetics-cell cycle model to analyse the response of human cells to intervention by topotecan, Computer Methods and Programs in Biomedicine 89 (2008) 169-178.

[108] Y.-H. Hsiang, M.G. Lihou, L.F. Liu, Arrest of replication forks by drug-stabilized topoisomerase I-DNA cleavable complexes as a mechanism of cell killing by camptothecin, Cancer Research 49 (1989) 5077-5082.

[109] D. Yang, J.T. Strode, H.P. Spielmann, A.H.J. Wang, T.G. Burke, DNA Interactions of Two Clinical Camptothecin Drugs Stabilize Their Active Lactone Forms, Journal of the American Chemical Society 120 (1998) 2979-2980.

[110] I. Husain, J.L. Mohler, H.F. Seigler, J.M. Besterman, Elevation of topoisomerase I messenger RNA, protein, and catalytic activity in human tumors: demonstration of tumor-type specificity and implications for cancer chemotherapy, Cancer Research 54 (1994) 539-546.

[111] S. Guichard, C. Terret, I. Hennebelle, I. Lochon, P. Chevreau, E. Frétigny, J. Selves, E. Chatelut, R. Bugat, P. Canal, CPT-11 converting carboxylesterase and topoisomerase I activities in tumour and normal colon and liver tissues, British Journal of Cancer 80 (1999) 364-370.

[112] H.L. McLeod, F. Douglas, M. Oates, R.P. Symonds, D. Prakash, A.G.J. Van Der Zee, S.B. Kaye, R. Brown, W.N. Keith, Topoisomerase I and II activity in human breast, cervix, lung and colon cancer, International Journal of Cancer 59 (1994) 607-611.

[113] S.H. Kaufmann, M. Charron, P.J. Burke, J.E. Karp, Changes in topoisomerase I levels and localization during myeloid maturation in vitro and in vivo, Cancer Research 55 (1995) 1255-1260.

[114] C.C. Shayo, A.G. Mladovan, A. Baldi, Differentiating agents modulate topoisomerase I activity in U-937 promonocytic cells, European Journal of Pharmacology 324 (1997) 129-133.

[115] Y.-P. Tsao, A. Russo, G. Nyamuswa, R. Silber, L.F. Liu, Interaction between replication forks and topoisomerase I-DNA cleavable complexes: studies in a cell-free SV40 DNA replication system, Cancer Research 53 (1993) 5908-5914.

[116] J. Wu, L.F. Liu, Processing of topoisomerase I cleavable complexes into DNA damage by transcription, Nucleic Acids Research 25 (1997) 4181-4186.

[117] C.J.H. Gerrits, M.J.A. de Jonge, J.H.M. Schellens, G. Stoter, J. Verweij, Topoisomerase I inhibitors: the relevance of prolonged exposure for present clinical development, British Journal of Cancer 76 (1997) 952-962.

[118] Q.C. Mao, J.D. Unadkat, Role of the breast cancer resistance protein (ABCG2) in drug transport, The AAPS Journal 7 (2005) E118-E133.

[119] B. Tan, D. Piwnica-Worms, L. Ratner, Multidrug resistance transporters and modulation, Current Opinion in Oncology 12 (2000) 450-458.

[120] S.V. Ambudkar, C. Kimchi-Sarfaty, Z.E. Sauna, M.M. Gottesman, P-glycoprotein: from genomics to mechanism, Oncogene 22 (2003) 7468-7485.

[121] C.B. Hendricks, E.K. Rowinsky, L.B. Grochow, R.C. Donehower, S.H. Kaufmann, Effect of P-glycoprotein expression on the accumulation and cytotoxicity of topotecan (SK&F 104864), a new camptothecin analogue, Cancer Research 52 (1992) 2268-2278.

[122] A. Haimeur, G. Conseil, R.G. Deeley, S.P. Cole, The MRP-related and BCRP/ABCG2 multidrug resistance proteins: biology, substrate specificity and regulation, Current Drug Metabolism 5 (2004) 21-53.

[123] L.A. Doyle, W. Yang, L.V. Abruzzo, T. Krogmann, Y. Gao, A.K. Rishi, D.D. Ross, A multidrug resistance transporter from human MCF-7 breast cancer cells, Proceedings of the National Academy of Sciences of the United States of America 95 (1998) 15665-15670.

[124] R. Allikmets, L.M. Schriml, A. Hutchinson, V. Romano-Spica, M. Dean, A human placenta-specific ATP-binding cassette gene (ABCP) on chromosome 4q22 that is involved in multidrug resistance, Cancer Research 58 (1998) 5337-5339.

[125] K. Miyake, L. Mickley, T. Litman, Z. Zhan, R. Robey, B. Cristensen, M. Brangi, L. Greenberger, M. Dean, T. Fojo, S.E. Bates, Molecular cloning of cDNAs which are highly overexpressed in mitoxantrone-resistant cells: demonstration of homology to ABC transport genes, Cancer Research 59 (1999) 8-13.

[126] E. Rocchi, A. Khodjakov, E.L. Volk, C.H. Yang, T. Litman, S.E. Bates, E. Schneider, The product of the ABC half-transporter gene ABCG2 (BCRP/MXR/ABCP) is expressed in the plasma membrane, Biochemical and Biophysical Research Communications 271 (2000) 42-46.

[127] M. Maliepaard, G.L. Scheffer, I.F. Faneyte, M.A. van Gastelen, A.C.L.M. Pijnenborg, A.H. Schinkel, M.J. van de Vijver, R.J. Scheper, J.H.M. Schellens,

Subcellular localization and distribution of the breast cancer resistance protein transporter in normal human tissues, Cancer Research 61 (2001) 3458-3464.

[128] J.D. Allen, A.H. Schinkel, Multidrug resistance and pharmacological protection mediated by the breast cancer resistance protein (BCRP/ABCG2), Molecular Cancer Therapeutics 1 (2002) 427-434.

[129] S.E. Bates, R. Robey, K. Miyake, K. Rao, D.D. Ross, T. Litman, The role of half-transporters in multidrug resistance, Journal of Bioenergetics and Biomembranes 33 (2001) 503-511.

[130] M. Maliepaard, M.A. van Gastelen, L.A. de Jong, D. Pluim, R.C.A.M. van Waardenburg, M.C. Ruevekamp-Helmers, B.G.J. Floot, J.H.M. Schellens, Overexpression of the BCRP/MXR/ABCP gene in a topotecan-selected ovarian tumor cell line, Cancer Research 59 (1999) 4559-4563.

[131] M. Ishii, M. Iwahana, I. Mitsui, M. Minami, S. Imagawa, A. Tohgo, A. Ejima, Growth inhibitory effect of a new camptothecin analog, DX-8951f, on various drug-resistant sublines including BCRP-mediated camptothecin derivative-resistant variants derived from the human lung cancer cell line PC-6, Anti-Cancer Drugs 11 (2000) 353-362.

[132] L.A. Doyle, D.D. Ross, Multidrug resistance mediated by the breast cancer resistance protein BCRP (ABCG2), Oncogene 22 (2003) 7340-7358.

[133] M. Dean, R. Allikmets, Complete characterization of the human ABC gene family, Journal of Bioenergetics and Biomembranes 33 (2001) 475-479.

[134] P. Borst, R.O. Elferink, Mammalian ABC transporters in health and disease, Annual Review of Biochemistry 71 (2002) 537-592.

[135] J.R. Riordan, Cystic fibrosis as a disease of misprocessing of the cystic fibrosis transmembrane conductance regulator glycoprotein, American Journal of Human Genetics 64 (1999) 1499-1504.

[136] J. Kartenbeck, U. Leuschner, R. Mayer, D. Keppler, Absence of the canalicular isoform of the MRP gene-encoded conjugate export pump from the hepatocytes in Dubin-Johnson syndrome, Hepatology 23 (1996) 1061-1066.

[137] H. Arnell, M. Mäntyjärvi, K. Tuppurainen, S. Andréasson, N. Dahl, Stargardt disease: linkage to the ABCR gene region on 1p21-p22 in Scandinavian families, Acta Ophthalmologica Scandinavica 76 (1998) 649-652.

[138] A.R. Walmsley, T.Q. Zhou, M.I. Borges-Walmsley, B.P. Rosen, A kinetic model for the action of a resistance efflux pump, Journal of Biological Chemistry 276 (2001) 6378-6391.

[139] D.D. Ross, W.D. Yang, L.V. Abruzzo, W.S. Dalton, E. Schneider, H. Lage, M. Dietel, L. Greenberger, S.P.C. Cole, L.A. Doyle, Atypical multidrug resistance: breast cancer resistance protein messenger RNA expression in mitroxantrone-selected cell lines, Journal of the National Cancer Institute 91 (1999) 429-433.

[140] G.A. Graf, L.Q. Yu, W.-P. Li, R. Gerard, P.L. Tuma, J.C. Cohen, H.H. Hobbs, ABCG5 and ABCG8 are obligate heterodimers for protein trafficking and biliary cholesterol excretion, Journal of Biological Chemistry 278 (2003) 48275-48282.

[141] J. Xu, Y. Liu, Y.Y. Yang, S. Bates, J.-T. Zhang, Characterization of oligomeric human half-ABC transporter ATP-binding cassette G2, Journal of Biological Chemistry 279 (2004) 19781-19789.

[142] C.H. Yang, E. Schneider, M.L. Kuo, E.L. Volk, E. Rocchi, Y.C. Chen, BCRP/MXR/ABCP expression in topotecan-resistant human breast carcinoma cells, Biochemical Pharmacology 60 (2000) 831-837.

[143] S.K. Rabindran, D.D. Ross, L.A. Doyle, W. Yang, L.M. Greenberger, Fumitremorgin C reverses multidrug resistance in cells transfected with the breast cancer resistance protein, Cancer Research 60 (2000) 47-50.

[144] R. Rajendra, M.K. Gounder, A. Saleem, J.H.M. Schellens, D.D. Ross, S.E. Bates, P. Sinko, E.H. Rubin, Differential effects of the breast cancer resistance protein on the cellular accumulation and cytotoxicity of 9-aminocamptothecin and 9-nitrocamptothecin, Cancer Research 63 (2003) 3228-3233.

[145] M. Yoshikawa, Y. Ikegami, S. Hayasaka, K. Ishi, A. Ito, K. Sano, T. Suzuki, T. Togawa, H. Yoshida, H. Soda, M. Oka, S. Kohno, S. Sawada, T. Ishikawa, S. Tanabe, Novel camptothecin analogues that circumvent ABCG2-associated drug resistance in human tumor cells, International Journal of Cancer 110 (2004) 921-927.

[146] A. Yoshida, A. Rzhetsky, L.C. Hsu, C. Chang, Human aldehyde dehydrogenase gene family, European Journal of Biochemistry 251 (1998) 549-557.

[147] R.A. Deltrich, Tissue and subcellular distribution of mammalian aldehyde-oxidizing capacity, Biochemical Pharmacology 15 (1966) 1911-1922.

[148] A. Yoshida, V. Davé, H. Han, K.J. Scanlon, Enhanced transcription of the cytosolic ALDH gene in cyclophosphamide resistant human carcinoma cells, Advances in Experimental Medicine and Biology 328 (1993) 63-72.

[149] J.S. Moreb, H.V. Baker, L.-J. Chang, M. Amaya, M.C. Lopez, B. Ostmark, W. Chou, ALDH isozymes downregulation affects cell growth, cell motility and gene expression in lung cancer cells, Molecular Cancer 7 (2008) DOI:10.1186/1476-4598-7-87.

[150] J.E. Visvader, G.J. Lindeman, Cancer stem cells in solid tumours: accumulating evidence and unresolved questions, Nature Reviews Cancer 8 (2008) 755-768.

[151] P.J. Smith, E. Furon, M. Wiltshire, L. Campbell, G.P. Feeney, R.D. Snyder, R.J. Errington, ABCG2-associated resistance to Hoechst 33342 and topotecan in a murine cell model with constitutive expression of side population characteristics, Cytometry Part A 75 (2009) 924-933.

[152] T. Litman, M. Brangi, E. Hudson, P. Fetsch, A. Abati, D.D. Ross, K. Miyake, J.H. Resau, S.E. Bates, The multidrug-resistant phenotype associated with overexpression of the new ABC half-transporter, MXR (ABCG2), Journal of Cell Science 113 (2000) 2011-2021.

[153] T. Hunter, J. Pines, Cyclins and cancer. II: Cyclin D and CDK inhibitors come of age, Cell 79 (1994) 573-582.

[154] B.E. Clurman, J.M. Roberts, Cell cycle and cancer, Journal of the National Cancer Institute 87 (1995) 1499-1501.

[155] P. Nurse, The incredible life and times of biological cells, Science 289 (2000) 1711-1716.

[156] K.A. Schafer, The cell cycle: a review, Veterinary Pathology 35 (1998) 461-478.

[157] J.J. Tyson, B. Novak, Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions, Journal of Theoretical Biology 210 (2001) 249-263.

[158] B. Novak, A. Csikasz-Nagy, B. Gyorffy, K. Chen, J.J. Tyson, Mathematical model of the fission yeast cell cycle with checkpoint controls at the G1/S, G2/M and metaphase/anaphase transitions, Biophysical Chemistry 72 (1998) 185-200.

[159] N. Thomas, I.D. Goodyer, Stealth sensors: real-time monitoring of the cell cycle, Targets 2 (2003) 26-33.

[160] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J.D. Watson, Molecular Biology of the Cell, Garland Publishing, New York, NY, USA, 1983.

[161] J.J.W. Baker, G.E. Allen, The Study of Biology, 3rd Ed., Addison-Wisley Publishing Company, Inc., Menlo Park, CA, USA, 1978.

[162] G.K. Chan, S.T. Liu, T.J. Yen, Kinetochore structure and function, Trends in Cell Biology 15 (2005) 589-598.

[163] A.W. Murray, T. Hunt, The Cell Cycle: An Introduction, W. H. Freeman & Co., New York, USA, 1993.

[164] S. Moreno, P. Nurse, Regulation of progression through the Gl phase of the cell cycle by the *rum1*$^{+}$ gene, Nature 367 (1994) 236-242.

[165] A.W. Murray, The genetics of cell cycle checkpoints, Current Opinion in Genetics & Development 5 (1995) 5-11.

[166] G.P. Feeney, R.J. Errington, M. Wiltshire, N. Marquez, S.C. Chappell, P.J. Smith, Tracking the cell cycle origins for escape from topotecan action by breast cancer cells, British Journal of Cancer 88 (2003) 1310-1317.

[167] T. Alarcón, R. Marches, K.M. Page, Mathematical models of the fate of lymphoma B cells after antigen receptor ligation with specific antibodies, Journal of Theoretical Biology 240 (2006) 54-71.

[168] J. Pines, Cyclins and cyclin-dependent kinases: theme and variations, Advances in Cancer Research 66 (1995) 181-212.

[169] L.L. Parker, H. Piwnica-Worms, Inactivation of the p34cdc2-cyclin B complex by the human WEE1 tyrosine kinase, Science 257 (1992) 1955-1957.

[170] S. van den Heuvel, E. Harlow, Distinct roles for cyclin-dependent kinases in cell cycle control, Science 262 (1993) 2050-2054.

[171] T. David-Pfeuty, Y. Nouvian-Dooghe, Human cyclin B1 is targeted to the nucleus in G1 phase prior to its accumulation in the cytoplasm, Oncogene 13 (1996) 1447-1460.

[172] R. Visintin, S. Prinz, A. Amon, CDC20 and CDH1: a family of substrate-specific activators of APC-dependent proteolysis, Science 278 (1997) 460-463.

[173] W. Zachariae, K. Nasmyth, Whose end is destruction: cell division and the anaphase-promoting complex, Genes & Development 13 (1999) 2039-2058.

[174] M.S. Greenblatt, W.P. Bennett, M. Hollstein, C.C. Harris, Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis, Cancer Research 54 (1994) 4855-4878.

[175] L.A. Donehower, M. Harvey, B.L. Slagle, M.J. McArthur, C.A. Montgomery Jr., J.S. Butel, A. Bradley, Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours, Nature 356 (1992) 215-221.

[176] C.J. Kemp, T. Wheldon, A. Balmain, p53-deficient mice are extremely susceptible to radiation-induced tumorigenesis, Nature Genetics 8 (1994) 66-69.

[177] A. Lavigueur, V. Maltby, D. Mock, J. Rossant, T. Pawson, A. Bernstein, High incidence of lung, bone, and lymphoid tumors in transgenic mice overexpressing mutant alleles of the p53 oncogene, Molecular and Cellular Biology 9 (1989) 3982-3991.

[178] V. Dulić, W.K. Kaufmann, S.J. Wilson, T.D. Tlsty, E. Lees, J.W. Harper, S.J. Elledge, S.I. Reed, p53-dependent inhibition of cyclin-dependent kinase activities in human fibroblasts during radiation-induced G1 arrest, Cell 76 (1994) 1013-1023.

[179] W.S. el-Deiry, J.W. Harper, P.M. O'Connor, V.E. Velculescu, C.E. Canman, J. Jackman, J.A. Pietenpol, M. Burrell, D.E. Hill, Y.S. Wang, K.G. Wiman, W.E. Mercer, M.B. Kastan, K.W. Kohn, S.J. Elledge, K.W. Kinzler, B. Vogelstein, WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis, Cancer Research 54 (1994) 1169-1174.

[180] O. Coqueret, New roles for p21 and p27 cell-cycle inhibitors: a function for each cell compartment?, Trends in Cell Biology 13 (2003) 65-70.

[181] J.W. Harper, G.R. Adami, N. Wei, K. Keyomarsi, S.J. Elledge, The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases, Cell 75 (1993) 805-816.

[182] S.M. Cross, C.A. Sanchez, C.A. Morgan, M.K. Schimke, S. Ramel, R.L. Idzerda, W.H. Raskind, B.J. Reid, A p53-dependent mouse spindle checkpoint, Science 267 (1995) 1353-1356.

[183] N. Stewart, G.G. Hicks, F. Paraskevas, M. Mowat, Evidence for a second cell cycle block at G2/M by p53, Oncogene 10 (1995) 109-115.

[184] M.L. Agarwal, A. Agarwal, W.R. Taylor, G.R. Stark, p53 controls both the G2/M and the G1 cell cycle checkpoints and mediates reversible growth arrest in human fibroblasts, Proceedings of the National Academy of Sciences of the United States of America 92 (1995) 8493-8497.

[185] M. Sugrue, C. Hutcheson, D. Fisk, C.G. Roberts, A. Mageed, J.R. Wingard, J.S. Moreb, The effect of overnight storage of leukapheresis stem cell products (LSCP) on cell viability, recovery, and cost., Journal of Hematotherapy 7 (1998) 431-436.

[186] S.A. Innocente, J.L.A. Abrahamson, J.P. Cogswell, J.M. Lee, p53 regulates a G2 checkpoint through cyclin B1, Proceedings of the National Academy of Sciences of the United States of America 96 (1999) 2147-2152.

[187] C. Cayrol, M. Knibiehler, B. Ducommun, p21 binding to PCNA causes G1 and G2 cell cycle arrest in p53-deficient cells, Oncogene 16 (1998) 311-320.

[188] R. Drapkin, G. Le Roy, H. Cho, S. Akoulitchev, D. Reinberg, Human cyclin-dependent kinase-activating kinase exists in three distinct complexes, Proceedings of the National Academy of Sciences of the United States of America 93 (1996) 6488-6493.

[189] A. Borgne, L. Meijer, Sequential dephosphorylation of p34(cdc2) on Thr-14 and Tyr-15 at the prophase/metaphase transition, Journal of Biological Chemistry 271 (1996) 27847-27854.

[190] R. Heald, M. McLoughlin, F. McKeon, Human wee1 maintains mitotic timing by protecting the nucleus from cytoplasmically activated Cdc2 kinase, Cell 74 (1993) 463-474.

[191] S. Kornbluth, B. Sebastian, T. Hunter, J. Newport, Membrane localization of the kinase which phosphorylates p34cdc2 on threonine 14, Molecular Biology of the Cell 5 (1994) 273-282.

[192] C.H. McGowan, P. Russell, Human Wee1 kinase inhibits cell division by phosphorylating p34cdc2 exclusively on Tyr15, The EMBO Journal 12 (1993) 75-85.

[193] C.H. McGowan, P. Russell, Cell cycle regulation of human WEE1, The EMBO Journal 14 (1995) 2166-2175.

[194] P.R. Mueller, T.R. Coleman, A. Kumagai, W.G. Dunphy, Myt1: a membrane-associated inhibitory kinase that phosphorylates Cdc2 on both threonine-14 and tyrosine-15, Science 270 (1995) 86-90.

[195] A. Devault, J.C. Cavadore, D. Fesquet, J.C. Labbé, T. Lorca, A. Picard, U. Strausfeld, M. Dorée, Concerted roles of cyclin A, cdc25+ mitotic inducer, and type 2A phosphatase in activating the cyclin B/cdc2 protein kinase at the G2/M phase transition, Cold Spring Harbor Symposia on Quantitative Biology 56 (1991) 503-513.

[196] R. Honda, Y. Ohba, A. Nagata, H. Okayama, H. Yasuda, Dephosphorylation of human p34cdc2 kinase on both Thr-14 and Tyr-15 by human cdc25B phosphatase, FEBS Letters 318 (1993) 331-334.

[197] B. Sebastian, A. Kakizuka, T. Hunter, Cdc25M2 activation of cyclin-dependent kinases by dephosphorylation of threonine-14 and tyrosine-15, Proceedings of the National Academy of Sciences of the United States of America 90 (1993) 3521-3524.

[198] E. Villa-Moruzzi, Activation of the cdc25C phosphatase in mitotic HeLa cells, Biochemical and Biophysical Research Communications 196 (1993) 1248-1254.

[199] I. Hoffmann, P.R. Clarke, M.J. Marcote, E. Karsenti, G. Draetta, Phosphorylation and activation of human cdc25-C by cdc2--cyclin B and its involvement in the self-amplification of MPF at mitosis, The EMBO Journal 12 (1993) 53-63.

[200] Z. Tang, T.R. Coleman, W.G. Dunphy, Two distinct mechanisms for negative regulation of the Wee1 protein kinase, The EMBO Journal 12 (1993) 3427-3436.

[201] N. Watanabe, M. Broome, T. Hunter, Regulation of the human WEE1Hu CDK tyrosine 15-kinase during the cell cycle, The EMBO Journal 14 (1995) 1878-1891.

[202] I. Chourpa, J.-M. Millot, G.D. Sockalingum, J.-F. Riou, M. Manfait, Kinetics of lactone hydrolysis in antitumor drugs of camptothecin series as studied by fluorescence spectroscopy, Biochimica et Biophysica Acta (BBA) - General Subjects 1379 (1998) 353-366.

[203] S.A. Streltsov, Action models for the antitumor drug camptothecin: formation of alkali-labile complex with DNA and inhibition of human DNA topoisomerase I, Journal of Biomolecular Structure & Dynamics 20 (2002) 447-454.

[204] A. Chen, V.T. Moy, Cross-linking of cell surface receptors enhances cooperativity of molecular adhesion, Biophysical Journal 78 (2000) 2814-2820.

[205] M.J. Chappell, Structural identifiability of models characterizing saturable binding: comparison of pseudo-steady-state and non-pseudo-steady-state model formulations, Mathematical Biosciences 133 (1996) 1-20.

[206] S. Wolfram, The Mathematica Book, 4th Ed., Mathematica Version 4, Wolfram Media/Cambridge University Press, Cambridge, UK, 1999.

[207] D. Bray, Protein molecules as computational elements in living cells, Nature 376 (1995) 307-312.

[208] N. Thomas, Lighting the circle of life: fluorescent sensors for covert surveillance of the cell cycle, Cell Cycle 2 (2003) 545-549.

[209] B. Neumann, M. Held, U. Liebel, H. Erfle, P. Rogers, R. Pepperkok, J. Ellenberg, High-throughput RNAi screening by time-lapse imaging of live human cells, Nature Methods 3 (2006) 385-390.

[210] K. Chu, N. Teele, M.W. Dewey, N. Albright, W.C. Dewey, Computerized video time lapse study of cell cycle delay and arrest, mitotic catastrophe, apoptosis and clonogenic survival in irradiated 14-3-3σ and CDKN1A (p21) knockout cell lines, Radiation Research 162 (2004) 270-286.

[211] I.A. Khan, P. Husemann, L. Campbell, N.S. White, R.J. White, P.J. Smith, R.J. Errington, ProgeniDB: a novel cell lineage database for generation associated phenotypic behavior in cell-based assays, Cell Cycle 6 (2007) 868-874.

[212] N. Marquez, S.C. Chappell, O.J. Sansom, A.R. Clarke, P. Teesdale-Spittle, R.J. Errington, P.J. Smith, Microtubule stress modifies intra-nuclear location of Msh2 in mouse embryonic fibroblasts, Cell Cycle 3 (2004) 662-671.

[213] I. Khan, J. Fisher, N. Marquez, P. Smith, R. Errington, A bioinformatics approach for the interrogation of molecular events in single cells: transforming fluorescent time-lapse microscopy images into numbers, BMC Systems Biology 1 (2007) Supplement 1: P31.

[214] K.W. Kohn, Molecular interaction map of the mammalian cell cycle control and DNA repair systems, Molecular Biology of the Cell 10 (1999) 2703-2734.

[215] T. Soussi, G. Lozano, p53 mutation heterogeneity in cancer, Biochemical and Biophysical Research Communications 331 (2005) 834-842.

[216] N.V. Guevara, H.S. Kim, E.I. Antonova, L. Chan, The absence of p53 accelerates atherosclerosis by increasing cell proliferation in vivo, Nature Medicine 5 (1999) 335-339.

[217] B.N. Goldstein, G. Ermakov, J.J. Centelles, H.V. Westerhoff, M. Cascante, What makes biochemical networks tick? A graphical tool for the identification of oscillophores, European Journal of Biochemistry 271 (2004) 3877-3887.

[218] C.M. Coelho, S.J. Leevers, Do growth and cell division rates determine cell size in multicellular organisms?, Journal of Cell Science 113 (2000) 2927-2934.

[219] T.T. Su, P.H. O'Farrell, Size control: cell proliferation does not equal growth, Current Biology 8 (1998) R687-R689.

[220] T. Hunt, Maturation promoting factor, cyclin and the control of M-phase, Current Opinion in Cell Biology 1 (1989) 268-274.

[221] B. Novak, J.J. Tyson, Numerical analysis of a comprehensive model of M-phase control in *Xenopus oocyte* extracts and intact embryos, Journal of Cell Science 106 (1993) 1153-1168.

[222] K.C. Chen, A. Csikasz-Nagy, B. Gyorffy, J. Val, B. Novak, J.J. Tyson, Kinetic analysis of a molecular model of the budding yeast cell cycle, Molecular Biology of the Cell 11 (2000) 369-391.

[223] V. Hatzimanikatis, K.H. Lee, J.E. Bailey, A mathematical description of regulation of the G1-S transition of the mammalian cell cycle, Biotechnology and Bioengineering 65 (1999) 631-637.

[224] A. Goldbeter, A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase, Proceedings of the National Academy of Sciences of the United States of America 88 (1991) 9107-9111.

[225] B.D. Aguda, A quantitative analysis of the kinetics of the G2 DNA damage checkpoint system, Proceedings of the National Academy of Sciences of the United States of America 96 (1999) 11352-11357.

[226] B.D. Aguda, Kick-starting the cell cycle: From growth-factor stimulation to initiation of DNA replication, Chaos 11 (2001) 269-276.

[227] B.D. Aguda, C.K. Algar, A structural analysis of the qualitative networks regulating the cell cycle and apoptosis, Cell Cycle 2 (2003) 538-544.

[228] B.D. Aguda, Y. Tang, The kinetic origins of the restriction point in the mammalian cell cycle, Cell Proliferation 32 (1999) 321-335.

[229] Z. Qu, J.N. Weiss, W.R. MacLellan, Regulation of the mammalian cell cycle: a model of the G1-to-S transition, American Journal of Physiology - Cell Physiology 284 (2003) C349-C364.

[230] Z. Qu, W.R. MacLellan, J.N. Weiss, Dynamics of the cell cycle: checkpoints, sizers, and timers, Biophysical Journal 85 (2003) 3600-3611.

[231] Z. Han, L. Yang, W.R. MacLellan, J.N. Weiss, Z. Qu, Hysteresis and cell cycle transitions: How crucial is it?, Biophysical Journal 88 (2005) 1626-1634.

[232] D. Angeli, J.E. Ferrell Jr., E.D. Sontag, Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feed back systems, Proceedings of the National Academy of Sciences of the United States of America 101 (2004) 1822-1827.

[233] S. Bai, D. Goodrich, C.D. Thron, E. Tecarro, M. Obeyesekere, Theoretical and experimental evidence for hysteresis in cell proliferation, Cell Cycle 2 (2003) 46-52.

[234] B. Novak, J.J. Tyson, Modeling the control of DNA replication in fission yeast, Proceedings of the National Academy of Sciences of the United States of America 94 (1997) 9147-9152.

[235] B. Novak, J.J. Tyson, A model for restriction point control of the mammalian cell cycle, Journal of Theoretical Biology 230 (2004) 563-579.

[236] C.D. Thron, Bistable biochemical switching and the control of the events of the cell cycle, Oncogene 15 (1997) 317-325.

[237] T.S. Gardner, M. Dolnik, J.J. Collins, A theory for controlling cell cycle dynamics using a reversibly binding inhibitor, Proceedings of the National Academy of Sciences of the United States of America 95 (1998) 14190-14195.

[238] M.N. Obeyesekere, S.L. Tucker, S.O. Zimmerman, A model for regulation of the cell cycle incorporating cyclin A, cyclin B and their complexes, Cell Proliferation 27 (1994) 105-113.

[239] M.N. Obeyesekere, E.S. Knudsen, J.Y.J. Wang, S.O. Zimmerman, A mathematical model of the regulation of the G1 phase of Rb+/+ and Rb-/- mouse embryonic fibroblasts and an osteosarcoma cell line, Cell Proliferation 30 (1997) 171-194.

[240] J.J. Tyson, B. Novak, K. Chen, J. Val, Checkpoints in the cell cycle from a modeler's perspective, Progress in Cell Cycle Research 1 (1995) 1-8.

[241] J.E. Ferrell Jr., Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability, Current Opinion in Cell Biology 14 (2002) 140-148.

[242] C. Cheng, L.M. Li, Systematic identification of cell cycle regulated transcription factors from microarray time series data, BMC Genomics 9 (2008) 116.

[243] O. Lavi, Y. Louzoun, What cycles the cell?–Robust autonomous cell cycle models, Mathematical Medicine and Biology 26 (2009) 337-359.

[244] J.R. Pomerening, S.Y. Kim, J.E. Ferrell Jr., Systems-level dissection of the cell-cycle oscillator: bypassing positive feedback produces damped oscillations, Cell 122 (2005) 565-578.

[245] R.G. Clyde, J.L. Bown, T.R. Hupp, N. Zhelev, J.W. Crawford, The role of modelling in identifying drug targets for diseases of the cell cycle, Journal of the Royal Society Interface 3 (2006) 617-627.

[246] P. Nurse, Universal control mechanism regulating onset of M-phase, Nature 344 (1990) 503-508.

[247] A. Csikász-Nagy, D. Battogtokh, K.C. Chen, B. Novák, J.J. Tyson, Analysis of a generic model of eukaryotic cell-cycle regulation, Biophysical Journal 90 (2006) 4361-4379.

[248] K. Hara, P. Tydeman, M. Kirschner, A cytoplasmic clock with the same period as the division cycle in Xenopus eggs, Proceedings of the National Academy of Sciences of the United States of America 77 (1980) 462-466.

[249] D.O. Morgan, Cyclin-dependent kinases: engines, clocks, and microprocessors, Annual Review of Cell and Developmental Biology 13 (1997) 261-291.

[250] A.W. Murray, M.W. Kirschner, Dominoes and clocks: the union of two views of the cell cycle, Science 246 (1989) 614-621.

[251] J.-M. Peters, J.R. Harris, D. Finley, (Eds.), Ubiquitin and the Biology of the Cell, Plenum Press, New York, 1998.

[252] A. Hershko, D. Ganoth, V. Sudakin, A. Dahan, L.H. Cohen, F.C. Luca, J.V. Ruderman, E. Eytan, Components of a system that ligates cyclin to ubiquitin and their regulation by the protein kinase cdc2, Journal of Biological Chemistry 269 (1994) 4940-4946.

[253] R.W. King, J.M. Peters, S. Tugendreich, M. Rolfe, P. Hieter, M.W. Kirschner, A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B, Cell 81 (1995) 279-288.

[254] V. Sudakin, D. Ganoth, A. Dahan, H. Heller, J. Hershko, F.C. Luca, J.V. Ruderman, A. Hershko, The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis, Molecular Biology of the Cell 6 (1995) 185-197.

[255] A.W. Murray, M.W. Kirschner, Cyclin synthesis drives the early embryonic cell cycle, Nature 339 (1989) 275-280.

[256] F.R. Cross, V. Archambault, M. Miller, M. Klovstad, Testing a mathematical model of the yeast cell cycle, Molecular Biology of the Cell 13 (2002) 52-70.

[257] C.D. Thron, A model for a bistable biochemical trigger of mitosis, Biophysical Chemistry 57 (1996) 239-251.

[258] D. McMillen, N. Kopell, J. Hasty, J.J. Collins, Synchronizing genetic relaxation oscillators by intercell signaling, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 679-684.

[259] J.M.G. Vilar, H.Y. Kueh, N. Barkai, S. Leibler, Mechanisms of noise-resistance in genetic oscillators, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 5988-5992.

[260] A. Goldbeter, Computational approaches to cellular rhythms, Nature 420 (2002) 238-245.

[261] P. Kaldis, M.J. Solomon, Analysis of CAK activities from human cells, European Journal of Biochemistry 267 (2000) 4213-4221.

[262] M.J. Solomon, J.W. Harper, J. Shuttleworth, CAK, the p34cdc2 activating kinase, contains a protein identical or closely related to p40MO15, The EMBO Journal 12 (1993) 3133-3142.

[263] P. Kaldis, The cdk-activating kinase (CAK): from yeast to mammals, Cellular and Molecular Life Sciences 55 (1999) 284-296.

[264] T. Lammers, S. Lavi, Role of type 2C protein phosphatases in growth regulation and in cellular stress signaling, Critical Reviews in Biochemistry and Molecular Biology 42 (2007) 437-461.

[265] D. Kültz, Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function, Journal of Experimental Biology 206 (2003) 3119-3124.

[266] J.P. Rogers, A.E. Beuscher, M. Flajolet, T. McAvoy, A.C. Nairn, A.J. Olson, P. Greengard, Discovery of protein phosphatase 2C inhibitors by virtual screening, Journal of Medicinal Chemistry 49 (2006) 1658-1667.

[267] G. Marlovits, C.J. Tyson, B. Novak, J.J. Tyson, Modeling M-phase control in Xenopus oocyte extracts: the surveillance mechanism for unreplicated DNA, Biophysical Chemistry 72 (1998) 169-184.

[268] G. Piaggio, A. Farina, D. Perrotti, I. Manni, P. Fuschi, A. Sacchi, C. Gaetano, Structure and growth-dependent regulation of the human cyclin B1 promoter, Experimental Cell Research 216 (1995) 396-402.

[269] A. Hagting, M. Jackman, K. Simpson, J. Pines, Translocation of cyclin B1 to the nucleus at prophase requires a phosphorylation-dependent nuclear import signal, Current Biology 9 (1999) 680-689.

[270] J. Yuan, F. Eckerdt, J. Bereiter-Hahn, E. Kurunci-Csacsko, M. Kaufmann, K. Strebhardt, Cooperative phosphorylation including the activity of polo-like kinase 1 regulates the subcellular localization of cyclin B1, Oncogene 21 (2002) 8282-8292.

[271] T. Alarcón, H.M. Byrne, P.K. Maini, A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells, Journal of Theoretical Biology 229 (2004) 395-411.

[272] V. Ramakrishnan, Histone structure and the organization of the nucleosome, Annual Review of Biophysics and Biomolecular Structure 26 (1997) 83-112.

[273] M.H.P. West, W.M. Bonner, Histone 2A, a heteromorphous family of eight protein species, Biochemistry 19 (1980) 3238-3245.

[274] T.H. Thatcher, M.A. Gorovsky, Phylogenetic analysis of the core histones H2A, H2B, H3, and H4, Nucleic Acids Research 22 (1994) 174-179.

[275] J.R. Pehrson, R.N. Fuji, Evolutionary conservation of histone macroH2A subtypes and domains, Nucleic Acids Research 26 (1998) 2837-2842.

[276] E.P. Rogakou, C. Boon, C. Redon, W.M. Bonner, Megabase chromatin domains involved in DNA double-strand breaks in vivo, Journal of Cell Biology 146 (1999) 905-916.

[277] T.T. Paull, E.P. Rogakou, V. Yamazaki, C.U. Kirchgessner, M. Gellert, W.M. Bonner, A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage, Current Biology 10 (2000) 886-895.

[278] T. Tanaka, X. Huang, H.D. Halicka, H. Zhao, F. Traganos, A.P. Albino, W. Dai, Z. Darzynkiewicz, Cytometry of ATM activation and histone H2AX phosphorylation to estimate extent of DNA damage induced by exogenous agents, Cytometry Part A 71 (2007) 648-661.

[279] C.J. Bakkenist, M.B. Kastan, DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation, Nature 421 (2003) 499-506.

[280] S. Burma, B.P. Chen, M. Murphy, A. Kurimasa, D.J. Chen, ATM phosphorylates histone H2AX in response to DNA double-strand breaks, Journal of Biological Chemistry 276 (2001) 42462-42467.

[281] X. Huang, F. Traganos, Z. Darzynkiewicz, DNA damage induced by DNA topoisomerase I- and topoisomerase II-inhibitors detected by histone H2AX phosphorylation in relation to the cell cycle phase and apoptosis, Cell Cycle 2 (2003) 614-619.

[282] O. Fernandez-Capetillo, A. Lee, M. Nussenzweig, A. Nussenzweig, H2AX: the histone guardian of the genome, DNA Repair 3 (2004) 959-967.

[283] M. Fragkos, J. Jurvansuu, P. Beard, H2AX is required for cell cycle arrest via the p53/p21 pathway, Molecular and Cellular Biology 29 (2009) 2828-2840.

[284] A.L. Gartel, S.K. Radhakrishnan, Lost in transcription: p21 repression, mechanisms, and consequences, Cancer Research 65 (2005) 3980-3985.

[285] G. Liu, G. Lozano, p21 stability: linking chaperones to a cell cycle checkpoint, Cancer Cell 7 (2005) 113-114.

[286] J.D. Friesner, B. Liu, K. Culligan, A.B. Britt, Ionizing radiation-dependent gamma-H2AX focus formation requires ataxia telangiectasia mutated and ataxia telangiectasia mutated and Rad3-related, Molecular Biology of the Cell 16 (2005) 2566-2576.

Appendix A

Structural identifiability analysis of the single-cell model

```
Lm'[t_] := -(Kom + Kmi) Lm[t] + Kcm Hm[t] + Kmo v0 Le[t]
Hm'[t_] := Kom Lm[t] - (Kcm + Kmi) Hm[t] + Kmo v0 He[t]
Le'[t_] := Kmi/v0 Lm[t] - (Kmo + Kom + Ki) Le[t] + Kcm He[t]
+ Ke/v1 Lc[t]
He'[t_] := Kmi/v0 Hm[t] + Kom Le[t] - (Kcm + Kmo) He[t] +
(Vm1 /v1) ((v3 Hc[t])/(v3 Km1 + Hc[t]))
Lc'[t_] := Ki v1 Le[t] - (Ke + Koc) Lc[t] + Kcc Hc[t] + Kdl
v2 Ln[t] - Kb (Bt - Ln[t]) Lc[t] - (Vm2 Lc[t])/(Km2 + Lc[t])
Hc'[t_] := Koc Lc[t] - Kcc Hc[t] + Kdh v2 Ln[t] - (Vm1 v3
Hc[t])/(v3 Km1 + Hc[t]) + (Vm2 Lc[t])/(Km2 + Lc[t])
Ln'[t_] := Kb/v2 (Bt - Ln[t]) Lc[t] - (Kdl + Kdh) Ln[t]
Lm[0]= (1+v0) d;Hm[0]=0;Le[0]=0;He[0]=0;
Lc[0]=0;Hc[0]=0;Ln[0]=0;

y[t_]:= {Le[t] + He[t],Lc[t] + Hc[t],Ln[t]}
v1 = (α v0)/(1+v0);
 y[0]
 {0,0,0}
 subst={Kmi→Kmib, Kmo→Kmob, Ki→Kib, Ke→Keb, Kb→Kbb,
Kdl→Kdlb, Kdh→Kdhb, Bt→Btb,v0→v0b,Vm1 →Vm1b, Km1→Km1b, Vm2
→Vm2b, Km2→Km2b}

{Kmi→Kmib,Kmo→Kmob,Ki→Kib,Ke→Keb,Kb→Kbb,Kdl→Kdlb,Kdh→Kdhb,B
t→Btb,v0→v0b,Vm1→Vm1b,Km1→Km1b,Vm2→Vm2b,Km2→Km2b}
 yDer=D[y[t], t];
yCoeff=yDer/.t→0
 {(d Kmi (1+v0))/v0,0,0}
 eqn=yCoeff[[1]]-(yCoeff[[1]]/.subst);
newSoln=Simplify[Solve[eqn==0, {Btb, v0b, Vm1b,Kmib,
Kmob,Kib,Keb, Kbb,Kdlb, Kdhb, Km1b, Vm2b, Km2b}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. ≫
 {{Kmib→(Kmi (1+v0) v0b)/(v0 (1+v0b))}}
 soln=newSoln[[1]]
 {Kmib→(Kmi (1+v0) v0b)/(v0 (1+v0b))}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 {-(d Kmi (Ki+Kmi+Kmo) (1+v0))/v0,d Ki Kmi α,0}

eqn=yCoeff-(yCoeff/.subst);
eqn=Simplify[eqn/.soln];
```

```
newSoln=Simplify[Solve[eqn==0,{Btb, v0b, Vm1b,Kmob,Kib,Keb,
Kbb,Kdlb, Kdhb, Km1b, Vm2b, Km2b}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. >>
 {{Kmob→(Ki (-v0+v0b))/((1+v0) v0b)+(Kmi v0+Kmo v0-Kmi
v0b+Kmo v0 v0b)/(v0+v0 v0b),Kib→(Ki v0 (1+v0b))/((1+v0)
v0b)}}
 soln=Simplify[soln/.newSoln[[1]]];
soln = Union [soln, newSoln [[1]]]
 {Kib→(Ki v0 (1+v0b))/((1+v0) v0b),Kmib→(Kmi (1+v0) v0b)/(v0
(1+v0b)),Kmob→(Ki (-v0+v0b))/((1+v0) v0b)+(Kmi v0+Kmo v0-Kmi
v0b+Kmo v0 v0b)/(v0+v0 v0b)}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 {(d Kmi (Ke Ki+Ki^2+(Kmi+Kmo)^2+Ki (Kmi+2 (Kmo+Kom)))
(1+v0))/v0,-d Ki Kmi (Bt Kb+Ke+Ki+Kmi+Kmo+2 Kom) α,(Bt d Kb
Ki Kmi α)/v2}

eqn=yCoeff-(yCoeff/.subst);
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{Btb, v0b, Vm1b,Keb, Kbb,Kdlb,
Kdhb, Km1b, Vm2b, Km2b}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. >>
 {{Btb→(Bt Kb)/Kbb,Keb→Ke,v0b→v0}}
 soln=Simplify[soln/.newSoln[[1]]];
soln = Union [soln, newSoln [[1]]]
 {Btb→(Bt Kb)/Kbb,Keb→Ke,Kib→Ki,Kmib→Kmi,Kmob→Kmo,v0b→v0}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 {-1/(Km1 Km2 v0) d Kmi (1+v0) (Bt Kb Ke Ki Km1 Km2+Ke^2 Ki
Km1 Km2+Ki^3 Km1 Km2+Ki^2 Km1 Km2 Kmi+Ki Km1 Km2 Kmi^2+Km1 Km2
Kmi^3+3 Ki^2 Km1 Km2 Kmo+4 Ki Km1 Km2 Kmi Kmo+3 Km1 Km2 Kmi^2
Kmo+3 Ki Km1 Km2 Kmo^2+3 Km1 Km2 Kmi Kmo^2+Km1 Km2 Kmo^3+3 Kcm Ki
Km1 Km2 Kom+3 Ki^2 Km1 Km2 Kom+3 Ki Km1 Km2 Kmi Kom+5 Ki Km1
Km2 Kmo Kom+3 Ki Km1 Km2 Kom^2-Ki Km2 Koc Vm1-Ki Vm1 Vm2+Ke Ki
Km1 (2 Ki Km2+Km2 (Kmi+2 Kmo+Koc+2 Kom)+Vm2)),1/(Km1 Km2) d
Ki Kmi (Bt^2 Kb^2 Km1 Km2+Ke^2 Km1 Km2+Ki^2 Km1 Km2+Ki Km1 Km2
Kmi+Km1 Km2 Kmi^2+2 Ki Km1 Km2 Kmo+2 Km1 Km2 Kmi Kmo+Km1 Km2
Kmo^2+3 Kcm Km1 Km2 Kom+3 Ki Km1 Km2 Kom+3 Km1 Km2 Kmi Kom+3
Km1 Km2 Kmo Kom+3 Km1 Km2 Kom^2-Km2 Koc Vm1-Vm1 Vm2+Bt Kb Km1
(Kdh Km2+Kdl Km2+2 Ke Km2+Ki Km2+Km2 Kmi+Km2 Kmo+Km2 Koc+2
Km2 Kom+Vm2)+Ke Km1 (2 Ki Km2+Km2 (Kmi+Kmo+Koc+2 Kom)+Vm2))
α,-(Bt d Kb Ki Kmi (Bt Kb Km2+Kdh Km2+Kdl Km2+Ke Km2+Ki
Km2+Km2 Kmi+Km2 Kmo+Km2 Koc+2 Km2 Kom+Vm2) α)/(Km2 v2)}

eqn=yCoeff-(yCoeff/.subst);
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{Vm1b, Kbb,Kdlb, Kdhb, Km1b,
Vm2b, Km2b}]];
 Solve::svars: Equations may not give solutions for all
```

```
"solve" variables. ≫
 newSoln[[1]]
 {Kdlb→Kdh-Kdhb+Kdl+Vm2/Km2-Vm2b/Km2b,Vm1b→(Km1b Km2 Km2b
Koc Vm1-Ke Km1 Km1b Km2b Vm2+Km1b Km2b Vm1 Vm2+Ke Km1 Km1b
Km2 Vm2b)/(Km1 Km2 Km2b Koc+Km1 Km2 Vm2b)}
 soln=Simplify[soln/.newSoln[[1]]];
soln = Union [soln, newSoln [[1]]]
 {Btb→(Bt Kb)/Kbb,Kdlb→Kdh-Kdhb+Kdl+Vm2/Km2-
Vm2b/Km2b,Keb→Ke,Kib→Ki,Kmib→Kmi,Kmob→Kmo,v0b→v0,Vm1b→(Km1b
Km2 Km2b Koc Vm1-Ke Km1 Km1b Km2b Vm2+Km1b Km2b Vm1 Vm2+Ke
Km1 Km1b Km2 Vm2b)/(Km1 Km2 Km2b Koc+Km1 Km2 Vm2b)}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0];

eqn=yCoeff-(yCoeff/.subst);
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{Kbb,Kdhb, Km1b, Vm2b,Km2b}]];
 Solve::svars: Equations may not give solutions for all
"solve" variables. ≫
 soln=Simplify[soln/.newSoln[[1]]];
soln = Union [soln, newSoln [[1]]]
 {Btb→(Bt
Kb)/Kbb,Kdhb→Kdh,Kdlb→Kdl,Keb→Ke,Kib→Ki,Kmib→Kmi,Kmob→Kmo,v
0b→v0,Vm1b→(Km1b Vm1)/Km1,Vm2b→(Km2b Vm2)/Km2}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0];
eqn=yCoeff-(yCoeff/.subst);
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{Kbb,Km1b, Km2b}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. ≫
 {{Kbb→Kb,Km2b→Km2}}
 soln=Simplify[soln/.newSoln[[1]]];
soln = Union [soln, newSoln [[1]]]

{Btb→Bt,Kbb→Kb,Kdhb→Kdh,Kdlb→Kdl,Keb→Ke,Kib→Ki,Km2b→Km2,Kmi
b→Kmi,Kmob→Kmo,v0b→v0,Vm1b→(Km1b Vm1)/Km1,Vm2b→Vm2}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0];
eqn=yCoeff-(yCoeff/.subst);
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{Km1b}]]
 {{Km1b→Km1}}
 soln=Simplify[soln/.newSoln[[1]]];
soln = Union [soln, newSoln [[1]]]

{Btb→Bt,Kbb→Kb,Kdhb→Kdh,Kdlb→Kdl,Keb→Ke,Kib→Ki,Km1b→Km1,Km2
b→Km2,Kmib→Kmi,Kmob→Kmo,v0b→v0,Vm1b→Vm1,Vm2b→Vm2}
```

Structural identifiability analysis of the extended cell cycle model

```
cyclin' [t_] := ksynth - kdest*(apcstar [t] + p21[t])*cyclin
[t] - ka*cdc2 [t]*cyclin [t] + kd*cdc2cyclin [t];

cdc2cyclin' [t_] := ka*cdc2 [t]*cyclin [t] - kd*cdc2cyclin
[t] - kdest*(apcstar [t] + p21[t])*cdc2cyclin [t] -
kwee1*wee1act [t]*cdc2cyclin [t]
- kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclin [t] +
kcdc25*cdc25act [t]*cdc2cyclinyp [t] +
kcdc25/factor*(C2*cdc25tot - cdc25act [t])*cdc2cyclinyp [t];

cdc2cyclinyp' [t_] := kwee1*wee1act [t]*cdc2cyclin
[t]+kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclin [t] -
kcdc25*cdc25act [t]*cdc2cyclinyp [t]
-kcdc25/factor*(C2*cdc25tot - cdc25act [t])*cdc2cyclinyp [t]
- kcak*cdc2cyclinyp [t] + kpp2c*cdc2cyclinyptp [t] -
kdest*(apcstar [t] + p21[t])*cdc2cyclinyp [t];

cdc2cyclinyptp'[t_] := kcak*cdc2cyclinyp [t] -
kpp2c*cdc2cyclinyptp [t] - kcdc25*cdc25act [t]*cdc2cyclinyptp
[t]
- kcdc25/factor*(C2*cdc25tot - cdc25act [t])*cdc2cyclinyptp
[t] + kwee1*wee1act [t]*cdc2cyclintp [t] +
kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclintp [t] -
kdest*(apcstar [t] + p21[t])*cdc2cyclinyptp [t];

cdc2cyclintp' [t_] := kcdc25*cdc25act [t]*cdc2cyclinyptp [t]
+ kcdc25/factor*(C2*cdc25tot - cdc25act[t])*cdc2cyclinyptp
[t]
- kwee1*wee1act [t]*cdc2cyclintp [t] -
kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclintp [t] -
kdest*(apcstar [t] + p21[t])*cdc2cyclintp [t];

cdc25act' [t_] := kcdc25on*(cdc2cyclintp
[t]^ncdc25/((C2*ec50cdc25)^ncdc25 + cdc2cyclintp
[t]^ncdc25))*(C2*cdc25tot - cdc25act [t]) -
kcdc25off*cdc25act [t];

wee1act' [t_] := - kwee1off*(cdc2cyclintp
[t]^nwee1/((C2*ec50wee1)^nwee1 + cdc2cyclintp
[t]^nwee1))*(wee1act [t])+kwee1on*(C2*wee1tot-wee1act [t]);

apc0' [t_] := - kapcon*(cdc2cyclintp
[t]^napc/((C2*ec50apc)^napc+cdc2cyclintp
[t]^napc))*(C2*apctot-apcstar [t])+kapcoff*apcstar [t];
```

```
p21'[t_] := kp21on*(cdc2cyclintp
[t]^np21/((C2*ec50p21)^np21+cdc2cyclintp
[t]^np21))*(C2*p21tot-p21 [t]) - kp21off*p21 [t];

apcstar [t_]:= C2*apctot-apc0 [t];

cdc2 [t_]:= C2*cdc2tot-(cdc2cyclin [t]+cdc2cyclinyp
[t]+cdc2cyclinyptp [t]+cdc2cyclintp [t]);

cyclin [0]= C2*0.141396; cdc2cyclin[0]= C2*0.992654;
cdc2cyclinyp[0]= C2*1.4763; cdc2cyclinyptp[0]=  C2*11.41023;
 cdc2cyclintp [0]= C2*1.64121; cdc25act[0]= C2*1.54009;
wee1act[0]= C2*15; p21[0] = C2*3.8539; apc0[0] = C2*46.1461;


nwee1=4;
ncdc25=4;
napc = 5;
np21=5;

y[t_]:= α cdc2cyclintp[t] + β;
 subst={α→αb,β→βb,C2→C2b, kp21off→kp21offb,p21tot→p21totb}
 {α→αb,β→βb,C2→C2b,kp21off→kp21offb,p21tot→p21totb}
 yCoeff=y[0]
 1.64121 C2 α+β
 eqn=yCoeff-(yCoeff/.subst)
newSoln=Simplify[Solve[eqn==0, {αb,C2b,βb,kp21offb,p21totb}]]
 1.64121 C2 α-1.64121 C2b αb+β-βb
 Solve::svars: Equations may not give solutions for all
"solve" variables. ≫
 {{βb→1.64121 C2 α-1.64121 C2b αb+1. β}}
 soln=newSoln[[1]]
 {βb→1.64121 C2 α-1.64121 C2b αb+1. β}
 yDer=D[y[t], t];
yCoeff=yDer/.t→0
 (17.5728 C2^2 kcdc25+(11.4102 C2 (-1.54009 C2+C2 cdc25tot)
kcdc25)/factor) α
 yCoeff
 2.52518 C2^2 α
 eqn=yCoeff-(yCoeff/.subst);
newSoln=Simplify[Solve[eqn==0, {αb,C2b,kp21offb,p21totb}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. ≫
 {{αb→(1. C2^2 α)/C2b^2}}
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 {βb→1.64121 C2 α-(1.64121 C2^2 α)/C2b+1. β}
 {αb→(C2^2 α)/C2b^2,βb→1.64121 C2 α-(1.64121 C2^2 α)/C2b+β}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 (0.107836 -0.388901 C2) C2^2 α
```

```
 eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{C2b,kp21offb,p21totb}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. >>
 {{C2b→0. +1. C2}}
 newSoln={{C2b→1.` C2}}
 {{C2b→1. C2}}
 (* First define the model' s parameters. Concentration units
are nM; time units are min

ksynth defines the constant rate of cyclin synthesis
and
kdest defines the rate constant for cyclin destruction. *)

ksynth=1.2;
kdest=0.005;

(* ka is the association constant for Cdc2+cyclin. kd is the
dissociation constant. kd/ka should be less than 1 nM *)

ka=0.1;
kd=0.001;

(* kwee1 and kcdc25 are the rate constants for the
inactivation and activation of Cdc2/cyclin by Wee1 and Cdc25.
If Km=1 uM and kcat=1 reaction/min,
then kcat/Km would be 1 uM-1 min-1 or 0.001 nM-1 min-1. The
bigger the value for 'factor', the stronger the feedback. If
factor=1 then the mitotic and interphase forms of Cdc25 and
Wee1 do not differ in activity,and so there is no feedback
from Cdc2 to Cdc25 and Wee1. If it's 10,the feedback is
strong enough to give sustained, realistic-looking
oscillations. *)

factor = 20;
kwee1=0.1;
kwee1basal=kwee1/factor;
kcdc25=0.1;
kcdc25basal=kcdc25/factor;

(* cdc2tot, cdc25tot, wee1tot and apctot are total
concentrations in nM *)

cdc2tot=100;
cdc25tot=15;
wee1tot=15;
apctot=50;


(* Assuming that the dependence of Cdc25 and Wee1 on Cdc2
```

```
activity is described by Hill functions with Hill
coefficients of nwee1 and ncdc25 and EC50's (in nM) of
ec50wee1 and ec50cdc25. *)

nwee1=4;
ncdc25=4;
napc=5;
np21 = 5;
ec50wee1=25;
ec50cdc25=25;
ec50apc=30;
ec50p21 = 30;

(* Define the rate constants for turning Cdc25 on (by Cdc2)
and off (by?) and for turning Wee1 off (by Cdc2) and on (by?)
*)

kcdc25on=0.8;
kcdc25off=0.08;
kapcon=0.8;
kapcoff=0.08;
kwee1on=0.8;
kwee1off=0.08;
kp21on = 0.8;

(* And define CAK activity and PP2C activity in nM-1 min-1 *)

kcak=0.8;
kpp2c=0.008;
 newSoln
 {{C2b→1. C2}}
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 {αb→1. α,βb→0. C2 α+β}
 {C2b→1. C2,αb→1. α,βb→0. C2 α+β}
 soln={C2b→ C2,αb→ α,βb→β}
 {C2b→C2,αb→α,βb→β}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 0.0598942 (-0.0627616+C2) C2^2 (4.45049 +C2) α
 eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{kp21offb,p21totb}]]
 {{}}
 soln
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 -0.00922425 (-0.0569411+C2) C2^2 (0.26181 +C2) (18.3566 +C2)
α
 eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
```

```
newSoln=Simplify[Solve[eqn==0,{kp21offb,p21totb}]]
 {{}}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 C2^2 (-0.000310113+2.6964 C2-2.50228 C2^3+0.00142062 C2^4+C2^2
(0.0542277 +0.00507981 kp21off-5.16716×10^-10 p21tot)) α

eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{kp21offb,p21totb}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. ≫
 {{kp21offb→kp21off-1.01719×10^-7 p21tot+1.01719×10^-7 p21totb}}
 {C2b→C2,αb→α,βb→β}
 {C2b→C2,kp21offb→kp21off-1.01719×10^-7 p21tot+1.01719×10^-7
p21totb,αb→α,βb→β}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 C2^2 (0.0000343614 -0.80468 C2+27.7678 C2^4-0.000218788 C2^5+C2^3
(0.769495 -0.00674463 kp21off-3.28906×10^-9 p21tot)+C2^2 (-
27.9778-0.00507981 kp21off^2+kp21off (-0.00139704+5.16716×10^-10
p21tot)+1.42106×10^-10 p21tot)) α
 eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{p21totb}]]
 {{p21totb→p21tot}}
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 {C2b→C2,kp21offb→kp21off+0. p21tot,αb→α,βb→β}
 {C2b→C2,kp21offb→kp21off,p21totb→p21tot,αb→α,βb→β}
```

Structural identifiability analysis of the coupled drug kinetics/dynamics model

Structural identifiability analysis of the drug kinetics model including the effect compartment

```
Lm'[t_] := -(Kom + Kmi) Lm[t] + Kcm Hm[t] + Kmo v0 Le[t]
Hm'[t_] := Kom Lm[t] - (Kcm + Kmi) Hm[t] + Kmo v0 He[t]
Le'[t_] := Kmi/v0 Lm[t] - (Kmo + Kom + Ki) Le[t] + Kcm He[t]
+ Ke/v1 Lc[t]
He'[t_] := Kmi/v0 Hm[t] + Kom Le[t] - (Kcm + Kmo) He[t] +
(Vm1 /v1) ((v3 Hc[t])/(v3 Km1 + Hc[t]))
```

```
Lc'[t_] := Ki v1 Le[t] - (Ke + Koc) Lc[t] + Kcc Hc[t] + Kdl
v2 Ln[t] - Kb (Bt - Ln[t]) Lc[t] - (Vm2 Lc[t])/(Km2 + Lc[t])
Hc'[t_] := Koc Lc[t] - Kcc Hc[t] + Kdh v2 Ln[t] - (Vm1 v3
Hc[t])/(v3 Km1 + Hc[t]) + (Vm2 Lc[t])/(Km2 + Lc[t])
Ln'[t_] := Kb/v2 (Bt - Ln[t]) Lc[t] - (Kdl + Kdh) Ln[t]

H2AX'[t_] := k11*Ln [t] - k2tp*H2AX [t];


y[t_]:= H2AX[t]

H2AX [0] = 0;

Lm[0]= (1+v0) d;Hm[0]=0;Le[0]=0;He[0]=0;
Lc[0]=0;Hc[0]=0;Ln[0]=0;

v1 = (α v0)/(1+v0);
 y[0]
 0
 subst={ k11→k11b, k2tp→k2tpb}
 {k11→k11b,k2tp→k2tpb}
 yDer=D[y[t], t]
yCoeff=yDer/.t→0
 -k2tp H2AX[t]+k11 Ln[t]
 0
 yDer=D[yDer, t]
yCoeff=Simplify[yDer/.t→0]
 -k2tp (-k2tp H2AX[t]+k11 Ln[t])+k11 ((Kb Lc[t] (Bt-
Ln[t]))/v2-(Kdh+Kdl) Ln[t])
 0
yDer=D[yDer, t]
yCoeff=Simplify[yDer/.t→0]
 -k2tp (-k2tp (-k2tp H2AX[t]+k11 Ln[t])+k11 ((Kb Lc[t] (Bt-
Ln[t]))/v2-(Kdh+Kdl) Ln[t]))+k11 (-(Kdh+Kdl) ((Kb Lc[t] (Bt-
Ln[t]))/v2-(Kdh+Kdl) Ln[t])-(Kb Lc[t] ((Kb Lc[t] (Bt-
Ln[t]))/v2-(Kdh+Kdl) Ln[t]))/v2+(Kb (Bt-Ln[t]) (Kcc Hc[t]-
(Ke+Koc) Lc[t]-(Vm2 Lc[t])/(Km2+Lc[t])+(Ki v0 α
Le[t])/(1+v0)-Kb Lc[t] (Bt-Ln[t])+Kdl v2 Ln[t]))/v2)
 0
 yDer=D[yDer, t]
yCoeff=Simplify[yDer/.t→0]
 (Bt d k11 Kb Ki Kmi α)/v2
 eqn=yCoeff-(yCoeff/.subst);
newSoln=Simplify[Solve[eqn==0, {k11b,k2tpb}]]
 Solve::svars: Equations may not give solutions for all
"solve" variables. ≫
 {{k11b→k11}}
 soln=newSoln[[1]]
 {k11b→k11}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0];
```

```
eqn=yCoeff-(yCoeff/.subst);
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{k2tpb}]]
 {{k2tpb→k2tp}}
 soln=Simplify[soln/.newSoln[[1]]];
soln = Union [soln, newSoln [[1]]]
 {k11b→k11,k2tpb→k2tp}
```

Structural identifiability analysis of the full drug kinetics/dynamics model

```
Lm'[t_] := -(Kom + Kmi) Lm[t] + Kcm Hm[t] + Kmo v0 Le[t]
Hm'[t_] := Kom Lm[t] - (Kcm + Kmi) Hm[t] + Kmo v0 He[t]
Le'[t_] := Kmi/v0 Lm[t] - (Kmo + Kom + Ki) Le[t] + Kcm He[t]
+ Ke/v1 Lc[t]
He'[t_] := Kmi/v0 Hm[t] + Kom Le[t] - (Kcm + Kmo) He[t] +
(Vm1 /v1) ((v3 Hc[t])/(v3 Km1 + Hc[t]))
Lc'[t_] := Ki v1 Le[t] - (Ke + Koc) Lc[t] + Kcc Hc[t] + Kdl
v2 Ln[t] - Kb (Bt - Ln[t]) Lc[t] - (Vm2 Lc[t])/(Km2 + Lc[t])
Hc'[t_] := Koc Lc[t] - Kcc Hc[t] + Kdh v2 Ln[t] - (Vm1 v3
Hc[t])/(v3 Km1 + Hc[t]) + (Vm2 Lc[t])/(Km2 + Lc[t])
Ln'[t_] := Kb/v2 (Bt - Ln[t]) Lc[t] - (Kdl + Kdh) Ln[t]

cyclin' [t_] := ksynth - kdest*(apcstar [t] + p21[t])*cyclin
[t] - ka*cdc2 [t]*cyclin [t] + kd*cdc2cyclin [t];

cdc2cyclin' [t_] := ka*cdc2 [t]*cyclin [t] - kd*cdc2cyclin
[t] - kdest*(apcstar [t] + p21[t])*cdc2cyclin [t] -
kwee1*wee1act [t]*cdc2cyclin [t]
- kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclin [t] +
kcdc25*cdc25act [t]*cdc2cyclinyp [t] +
kcdc25/factor*(C2*cdc25tot - cdc25act [t])*cdc2cyclinyp [t];

cdc2cyclinyp' [t_] := kwee1*wee1act [t]*cdc2cyclin
[t]+kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclin [t] -
kcdc25*cdc25act [t]*cdc2cyclinyp [t]
-kcdc25/factor*(C2*cdc25tot - cdc25act [t])*cdc2cyclinyp [t]
- kcak*cdc2cyclinyp [t] + kpp2c*cdc2cyclinyptp [t] -
kdest*(apcstar [t] + p21[t])*cdc2cyclinyp [t];

cdc2cyclinyptp'[t_] := kcak*cdc2cyclinyp [t] -
kpp2c*cdc2cyclinyptp [t] - kcdc25*cdc25act [t]*cdc2cyclinyptp
[t]
- kcdc25/factor*(C2*cdc25tot - cdc25act [t])*cdc2cyclinyptp
[t] + kwee1*wee1act [t]*cdc2cyclintp [t] +
kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclintp [t] -
kdest*(apcstar [t] + p21[t])*cdc2cyclinyptp [t];
```

```
cdc2cyclintp' [t_] := kcdc25*cdc25act [t]*cdc2cyclinyptp [t]
+ kcdc25/factor*(C2*cdc25tot - cdc25act[t])*cdc2cyclinyptp
[t]
- kwee1*wee1act [t]*cdc2cyclintp [t] -
kwee1/factor*(C2*wee1tot - wee1act [t])*cdc2cyclintp [t] -
kdest*(apcstar [t] + p21[t])*cdc2cyclintp [t];

cdc25act' [t_] := kcdc25on*(cdc2cyclintp
[t]^ncdc25/((C2*ec50cdc25)^ncdc25 + cdc2cyclintp
[t]^ncdc25))*(C2*cdc25tot - cdc25act [t]) -
kcdc25off*cdc25act [t];

wee1act' [t_] := - kwee1off*(cdc2cyclintp
[t]^nwee1/((C2*ec50wee1)^nwee1 + cdc2cyclintp
[t]^nwee1))*(wee1act [t])+kwee1on*(C2*wee1tot-wee1act [t]);

apc0' [t_] := - kapcon*(cdc2cyclintp
[t]^napc/((C2*ec50apc)^napc+cdc2cyclintp
[t]^napc))*(C2*apctot-apcstar [t])+kapcoff*apcstar [t];

p21'[t_] := kp21on*(cdc2cyclintp
[t]^np21/((C2*ec50p21)^np21+cdc2cyclintp
[t]^np21))*(C2*p21tot-p21 [t]) - (kp21off*p21 [t] )/(1
+k2tp*H2AX[t]);

H2AX'[t_] := k11*Ln [t] - k2tp*H2AX [t];

apcstar [t_]:= C2*apctot-apc0 [t];

cdc2 [t_]:= C2*cdc2tot-(cdc2cyclin [t]+cdc2cyclinyp
[t]+cdc2cyclinyptp [t]+cdc2cyclintp [t]);

y[t_]:= {α cdc2cyclintp[t] + β}

cyclin [0]= C2*0.141396; cdc2cyclin[0]= C2*0.992654;
cdc2cyclinyp[0]= C2*1.4763; cdc2cyclinyptp[0]=  C2*11.41023;
cdc2cyclintp [0]= C2*1.64121; cdc25act[0]= C2*1.54009;
wee1act[0]= C2*15; p21[0] = C2*3.8539; apc0[0] = C2*46.1461;

H2AX [0] = a8;

Lm[0]= a1;Hm[0]=a2;Le[0]=a3;He[0]=a4;
Lc[0]=a5;Hc[0]=a6;Ln[0]=a7;

v1 = (α1 v0)/(1+v0);
 subst={α→αb,β→βb,C2→C2b, kp21off→kp21offb,p21tot→p21totb}
 {α→αb,β→βb,C2→C2b,kp21off→kp21offb,p21tot→p21totb}
 nwee1=4;
ncdc25=4;
napc = 5;
```

```
np21=5;
 yCoeff=y[0]
 {1.64121 C2 α+β}
 eqn=yCoeff-(yCoeff/.subst)
newSoln=Simplify[Solve[eqn==0, {αb,C2b,βb,kp21offb,p21totb}]]
 {1.64121 C2 α-1.64121 C2b αb+β-βb}
 Solve::svars :  Equations may not give solutions for all
"solve" variables. ≫
 {{βb→1.64121 C2 α-1.64121 C2b αb+1. β}}
 soln=newSoln[[1]]
 {βb→1.64121 C2 α-1.64121 C2b αb+1. β}
 yDer=D[y[t], t];
yCoeff=yDer/.t→0
 {(17.5728 C2^2 kcdc25+(11.4102 C2 (-1.54009 C2+C2 cdc25tot)
kcdc25)/factor) α}
 yCoeff
 eqn=yCoeff-(yCoeff/.subst);
newSoln=Simplify[Solve[eqn==0, {αb,C2b,kp21offb,p21totb}]]
 Solve::svars :  Equations may not give solutions for all
"solve" variables. ≫
 {{αb→(C2^2 α)/C2b^2}}
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 {βb→1.64121 C2 α-(1.64121 C2^2 α)/C2b+1. β}
 {αb→(C2^2 α)/C2b^2,βb→1.64121 C2 α-(1.64121 C2^2 α)/C2b+1. β}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 {C2^2 kcdc25 (2.27363 kcak-27.0637 C2 kcdc25+(127.496
kcdc25off+17.5728 ec50cdc25^4 kcdc25off+127.496 kcdc25on-
82.7849 cdc25tot kcdc25on)/(7.25532 factor+ec50cdc25^4
factor)+((-127.496-17.5728 ec50cdc25^4) kcdc25off+(-
127.496+82.7849 cdc25tot) kcdc25on)/(7.25532 +ec50cdc25^4)+((-
1.54009+cdc25tot) (1.4763 kcak-17.5728 C2 kcdc25-11.4102
kpp2c))/factor-17.5728 kpp2c) α}

eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{C2b,kp21offb,p21totb}]]
 Solve::svars :  Equations may not give solutions for all
"solve" variables. ≫
 {{C2b→0. +1. C2}}
 newSoln
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 {αb→(C2^2 α)/(0. +1. C2)^2,βb→(1. C2 β)/(0. +1. C2)}
 {C2b→C2,αb→α,βb→β}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 {0.0598942 (-0.0627616+C2) C2^2 (4.45049 +C2) α}
 eqn=Simplify[yCoeff-(yCoeff/.subst)]
```

```
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{kp21offb,p21totb}]]
 {0.0598942 (-0.0627616+C2) C2^2 (4.45049 +C2) α-0.0598942 (-
0.0627616+C2b) C2b^2 (4.45049 +C2b) αb}
 {{}}
 eqn
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 {-0.00922425 (-0.0569411+C2) C2^2 (0.26181 +C2) (18.3566 +C2)
α}

eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{kp21offb,p21totb}]]
 {{}}
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 (* First define the model' s parameters. Concentration units
are nM; time units are min

ksynth defines the constant rate of cyclin synthesis
and
kdest defines the rate constant for cyclin destruction. *)

ksynth=1.2;
kdest=0.005;

(* ka is the association constant for Cdc2+cyclin. kd is the
dissociation constant. kd/ka should be less than 1 nM *)

ka=0.1;
kd=0.001;

(* kwee1 and kcdc25 are the rate constants for the
inactivation and activation of Cdc2/cyclin by Wee1 and Cdc25.
If Km=1 uM and kcat=1 reaction/min,
then kcat/Km would be 1 uM-1 min-1 or 0.001 nM-1 min-1. The
bigger the value for 'factor', the stronger the feedback. If
factor=1 then the mitotic and interphase forms of Cdc25 and
Wee1 do not differ in activity,and so there is no feedback
from Cdc2 to Cdc25 and Wee1. If it's 10,the feedback is
strong enough to give sustained, realistic-looking
oscillations. *)

factor = 20;
kwee1=0.1;
kwee1basal=kwee1/factor;
kcdc25=0.1;
kcdc25basal=kcdc25/factor;

(* cdc2tot, cdc25tot, wee1tot and apctot are total
```

```
concentrations in nM *)

cdc2tot=100;
cdc25tot=15;
wee1tot=15;
apctot=50;


(* Assuming that the dependence of Cdc25 and Wee1 on Cdc2
activity is described by Hill functions with Hill
coefficients of nwee1 and ncdc25 and EC50's (in nM) of
ec50wee1 and ec50cdc25. *)

nwee1=4;
ncdc25=4;
napc=5;
np21 = 5;
ec50wee1=25;
ec50cdc25=25;
ec50apc=30;
ec50p21 = 30;

(* Define the rate constants for turning Cdc25 on (by Cdc2)
and off (by?) and for turning Wee1 off (by Cdc2) and on (by?)
*)

kcdc25on=0.8;
kcdc25off=0.08;
kapcon=0.8;
kapcoff=0.08;
kwee1on=0.8;
kwee1off=0.08;
kp21on = 0.8;

(* And define CAK activity and PP2C activity in nM-1 min-1 *)

kcak=0.8;
kpp2c=0.008;
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0]
 {1/(1. +a8 k2tp) C2^2 (-0.000310113-0.000310113 a8 k2tp+C2^3
(-2.50228-2.50228 a8 k2tp)+C2^4 (0.00142062 +0.00142062 a8
k2tp)+C2 (2.6964 +2.6964 a8 k2tp)+C2^2 (0.0542277 +0.00507981
kp21off+a8 k2tp (0.0542277 -5.16716×10^-10 p21tot)-5.16716×10^-10
p21tot)) α}

eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{kp21offb,p21totb}]]
 Solve::svars :  Equations may not give solutions for all
"solve" variables. ≫
```

```
 {{kp21offb→(1.01719×10^-7+1.01719×10^-7 a8 k2tp) p21totb+1/(C2^4
α) 196.858 (1. +a8 k2tp) ((0. C2^4 α)/(1. +a8 k2tp)+(0. a8 C2^4
k2tp α)/(1. +a8 k2tp)+(C2^2 (0. +0. a8 k2tp) α)/(1. +a8
k2tp)+(C2^3 (0. +0. a8 k2tp) α)/(1. +a8 k2tp)+(C2^5 (0. +0. a8
k2tp) α)/(1. +a8 k2tp)+(C2^6 (0. +0. a8 k2tp) α)/(1. +a8
k2tp)+(0.00507981 C2^4 kp21off α)/(1. +a8 k2tp)-(5.16716×10^-10
C2^4 p21tot α)/(1. +a8 k2tp)-(5.16716×10^-10 a8 C2^4 k2tp p21tot
α)/(1. +a8 k2tp))}}
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 {C2b→C2,αb→α,βb→β}
 {C2b→C2,kp21offb→(1.01719×10^-7+1.01719×10^-7 a8 k2tp)
p21totb+1/(C2^4 α) 196.858 (1. +a8 k2tp) ((0. C2^4 α)/(1. +a8
k2tp)+(0. a8 C2^4 k2tp α)/(1. +a8 k2tp)+(C2^2 (0. +0. a8 k2tp)
α)/(1. +a8 k2tp)+(C2^3 (0. +0. a8 k2tp) α)/(1. +a8 k2tp)+(C2^5
(0. +0. a8 k2tp) α)/(1. +a8 k2tp)+(C2^6 (0. +0. a8 k2tp)
α)/(1. +a8 k2tp)+(0.00507981 C2^4 kp21off α)/(1. +a8 k2tp)-
(5.16716×10^-10 C2^4 p21tot α)/(1. +a8 k2tp)-(5.16716×10^-10 a8
C2^4 k2tp p21tot α)/(1. +a8 k2tp)),αb→α,βb→β}
 yDer=D[yDer, t];
yCoeff=Simplify[yDer/.t→0];
eqn=Simplify[yCoeff-(yCoeff/.subst)];
eqn=Simplify[eqn/.soln];
newSoln=Simplify[Solve[eqn==0,{p21totb}]]
 {{p21totb→p21tot}}
 newSoln={{p21totb→p21tot}}
 soln=Simplify[soln/.newSoln[[1]]]
soln=Union[soln,newSoln[[1]]]
 {C2b→C2,kp21offb→(1.01719×10^-7+1.01719×10^-7 a8 k2tp)
p21tot+1/(C2^4 α) 196.858 (1. +a8 k2tp) ((0. C2^4 α)/(1. +a8
k2tp)+(0. a8 C2^4 k2tp α)/(1. +a8 k2tp)+(C2^2 (0. +0. a8 k2tp)
α)/(1. +a8 k2tp)+(C2^3 (0. +0. a8 k2tp) α)/(1. +a8 k2tp)+(C2^5
(0. +0. a8 k2tp) α)/(1. +a8 k2tp)+(C2^6 (0. +0. a8 k2tp)
α)/(1. +a8 k2tp)+(0.00507981 C2^4 kp21off α)/(1. +a8 k2tp)-
(5.16716×10^-10 C2^4 p21tot α)/(1. +a8 k2tp)-(5.16716×10^-10 a8
C2^4 k2tp p21tot α)/(1. +a8 k2tp)),αb→α,βb→β}
 {C2b→C2,k11b→k11,k2tpb→1.
k2tp,kp21offb→kp21off,p21totb→p21tot,αb→α,βb→β}
```

Appendix B

Parameter estimation for the hydrolysis model

```
* TPT in Buffer Parameter Estimation pH = 7.2 Sept. 2010   ;
*========================================================= ;
* Mohammed Isam Atari                                      ;
*========================================================= ;

EXECUTE ;
OPEN 8 "Pstream1In.txt" ;
OPEN 9 "Pstream2In.txt" ;
OPEN 10 "Plot1.out";
OPEN 11 "Plot2.out";
OPEN 12 "Fit_Plot1_Data.txt";
OPEN 13 "Fit_Plot2_Data.txt";
**;

VARIABLE
x1a x2a x1b x2b
;

* Define the Parameters ;
PARAMETER
k12 1.8e-2 k21 9.36e-3 D 10
;

INTEGER #COUNT ;


* Assign intital conditions for variables ;
COMPILE INITIAL;
x1a = D ;
x2a = 0 ;
x1b = D ;
x2b = 0 ;
**;

COMPILE EQUATIONS ;
'x1a = -k21*x1a + k12*x2a ;
'x2a = k21*x1a - k12*x2a ;
'x1b = -k21*x1b + k12*x2b ;
'x2b = k21*x1b - k12*x2b ;
**;

SETPSTREAM 1 8;
**;
SETPSTREAM 2 9;
**;


* DATA relative error estimate default 0.05;
DATA 0.05 ;
TIME x1a x2a x1b x2b ;
RANGE 9.99 3.11 9.93 3.12;
1.00  9.99  0.00  9.93  0.05  ;
```

```
2.00  9.89  0.09  9.85  0.11  ;
5.00  9.71  0.27  9.63  0.29  ;
10.00 9.36  0.60  9.26  0.71  ;
15.00 8.98  1.10  8.91  1.02  ;
30.00 7.96  2.01  8.40  1.52  ;
45.00 7.47  2.48  7.34  2.61  ;
60.00 7.10  3.07  6.86  3.08  ;
120.00      6.79  3.11  6.83  3.12  ;
**;

* Identify parameters to be varied;
SETVARY k12 k21 ;

BEGIN;

COMPILE INSTANT ;
WRITE 1=10, "PRINT STREAM NO." % ;
WRITE 1, "TIME La La(obs) Ha Ha(obs) " % ;
WRITE 1=11, "PRINT STREAM NO." % ;
WRITE 1, "TIME Lb Lb(obs) Hb Hb(obs) " % ;
WRITE 1=12, "PRINT STREAM NO." % ;
WRITE 1, "TIME La La(obs) Ha Ha(obs) " % ;
WRITE 1=13, "PRINT STREAM NO." % ;
WRITE 1, "TIME Lb Lb(obs) Hb Hb(obs) " % ;
#COUNT=0;
**;

COMPILE PRINT;
* Output routine called during the final phase;
#COUNT=#COUNT+1;
DO 10 FOR #2=#COUNT-1 ;
WRITE 1=10, ((E14,6)) TIME, x1a, VOBS<0,#2>, x2a, VOBS<1,#2> ;
WRITE 1=11, ((E14,6)) TIME, x1b, VOBS<2,#2>, x2b, VOBS<3,#2> ;
WRITE 1=12, ((E14,6)) TIME, x1a, VOBS<0,#2>, x2a, VOBS<1,#2> ;
WRITE 1=13, ((E14,6)) TIME, x1b, VOBS<2,#2>, x2b, VOBS<3,#2> ;
LABEL 10 ;
**;

SETNOFIT;
WHENEVER TIME = TOBS % CALL PRINT;
**;

BEGIN;
STOP;
```

Parameter estimation of the single-cell model

Averaged data

```
* Active Pump Parameter Fitting
*=========================================================== ;
* Mohammed Isam Atari                                        ;
*=========================================================== ;
```

```
EXECUTE ;

OPEN 8 "Pstream1In.txt";
OPEN 9 "Pstream2In.txt";
OPEN 10 "fplot.out";
OPEN 11 "fit_plot_data.txt";
**;

* Define the Variables ;
VARIABLE Lm Hm Le He Lc Hc Ln;

* Define the Parameters ;
PARAMETER
ki ke kb kmi kmo kom kcm koc kcc kdl kdh
v0 v1 v2 v3
D
Vm1 Km1 Vm2 Km2
alpha 2.4375e4
BT Med Cyt Nuc
;

* Open file, read in and assign parameters ;
COMPILE INSTANT;
OPEN 7 "pars.txt";
READ 7 kom ;
READ 7 kcm ;
READ 7 koc ;
READ 7 kcc ;
READ 7 ki ;
READ 7 ke ;
READ 7 kb ;
READ 7 kmi ;
READ 7 kmo ;
READ 7 kdl ;
READ 7 kdh ;
READ 7 v0 ;
READ 7 v2 ;
READ 7 v3 ;
READ 7 D ;
READ 7 Vm1 ;
READ 7 Km1 ;
READ 7 Vm2 ;
READ 7 Km2 ;
READ 7 BT ;
**;

COMPILE GENERAL ;
Med = Le + He ;
Nuc = Ln ;
Cyt = Lc + Hc ;
**;


INTEGER #COUNT ;

* Assign Intital Conditions for Variables ;
COMPILE INITIAL;
Lm = (1+v0)*D ;
Hm = 0 ;
Le = 0 ;
```

```
He = 0 ;
Lc = 0 ;
Hc = 0 ;
Ln = 0 ;
v1 = alpha*v0/(1+v0) ;
**;

COMPILE EQUATIONS;
'Lm = -(kom + kmi)*Lm + kcm*Hm + kmo*v0*Le ;
'Hm = kom*Lm -(kcm + kmi)*Hm + kmo*v0*He ;
'Le = (kmi*Lm)/v0 -(kmo + kom + ki)*Le + kcm*He +
(ke*Lc)/v1 ;
'He = (kmi*Hm)/v0 + kom*Le -(kcm + kmo)*He +
(Vm1*v3*Hc)/((v1)*(v3*Km1 + Hc)) ;
'Lc = ki*v1*Le -(ke + koc)*Lc + kcc*Hc + kdl*v2*Ln - kb*(BT - Ln)*Lc
- (Vm2*Lc)/(Km2 + Lc) ;
'Hc = koc*Lc - kcc*Hc + kdh*v2*Ln - (v3*Vm1*Hc)/(v3*Km1 + Hc)
+ (Vm2*Lc)/(Km2 + Lc) ;
'Ln = kb*Lc*(BT - Ln)/v2 - (kdl + kdh)*Ln;
**;


SETPSTREAM 1 8;
**;
SETPSTREAM 2 9;
**;

DATA 0.05;
TIME Med Nuc Cyt ;
RANGE  10.7556 12.0667 9.2068 ;
5     3.4000      0.0632      0.1538 ;
10    5.8889      0.1932      0.2940 ;
15    6.9778      0.3453      0.4701 ;
20    7.5333      0.5385      0.6547 ;
25    7.9111      0.7915      0.8325 ;
30    8.2889      1.0496      1.0667 ;
35    8.5778      1.3265      1.3009 ;
40    8.7111      1.6188      1.5162 ;
45    8.8667      1.8684      1.7385 ;
50    9.0000      2.1641      1.9658 ;
55    9           2.4085      2.1863 ;
60    9.0222      2.6581      2.3829 ;
65    9.1778      2.9368      2.5590 ;
70    9.2222      3.1658      2.6889 ;
75    9.2         3.3658      2.9453 ;
80    9.2222      3.6137      3.0838 ;
85    9.2889      3.8291      3.2496 ;
90    9.3111      4.0547      3.4291 ;
95    9.3556      4.2462      3.5641 ;
100   9.4444      4.4479      3.6991 ;
105   9.3778      4.6222      3.8291 ;
110   9.4         4.8325      4.0291 ;
115   9.4444      4.9983      4.0547 ;
120   9.4222      5.147 4.1761 ;
125   9.4         5.3368      4.2034 ;
130   9.4444      5.4701      4.3726 ;
135   9.4667      5.653 4.4718 ;
140   9.5556      5.7658      4.6325 ;
145   9.5333      5.9487      4.6786 ;
150   9.6         6.1504      4.7932 ;
155   9.5556      6.2547      4.8427 ;
```

```
160   9.6222      6.4496      4.9778 ;
165   9.7333      6.6222      5.0855 ;
170   9.7333      6.7094      5.2547 ;
175   9.8         6.8803      5.3128 ;
180   9.7778      7.0718      5.4821 ;
185   9.7778      7.1744      5.4872 ;
190   9.8667      7.3658      5.6256 ;
195   9.8667      7.465 5.7060 ;
200   9.7778      7.5658      5.7573 ;
205   9.7333      7.665 5.8256 ;
210   9.7556      7.7641      6.0239 ;
215   9.7556      7.935 5.9726 ;
220   9.8444      8.0581      6.0821 ;
225   9.8222      8.2359      6.1932 ;
230   9.7556      8.3248      6.2701 ;
235   9.8         8.4547      6.4222 ;
240   9.8889      8.5556      6.4667 ;
245   9.8444      8.6684      6.5675 ;
250   9.9111      8.7538      6.6376 ;
255   9.8889      8.8735      6.6957 ;
260   9.9333      8.9778      6.7538 ;
265   9.9556      9.0974      6.9060 ;
270   10          9.1966      6.9880 ;
275   10          9.3863      7.0889 ;
280   10.0222     9.4291      7.1111 ;
285   10.0222     9.5385      7.1658 ;
290   10.0889     9.6735      7.3282 ;
295   10.1556     9.7641      7.3077 ;
300   10.1333     9.8821      7.4222 ;
305   10.1778     9.9504      7.4803 ;
310   10.1778     10.0256     7.5470 ;
315   10.2889     10.1368     7.6650 ;
320   10.2222     10.1709     7.6906 ;
325   10.2222     10.2462     7.7521 ;
330   10.2444     10.3658     7.7179 ;
335   10.2222     10.453      7.8701 ;
340   10.2889     10.5197     7.9402 ;
345   10.2889     10.641      8.0137 ;
350   10.3333     10.8154     8.0803 ;
355   10.3111     10.788      8.0838 ;
360   10.3333     10.8427     8.1470 ;
365   10.3778     10.9402     8.2889 ;
370   10.4222     11.0547     8.2530 ;
375   10.4222     11.0496     8.3504 ;
380   10.3556     11.1145     8.3385 ;
385   10.4222     11.1556     8.4017 ;
390   10.4        11.241      8.5026 ;
395   10.3778     11.2342     8.6120 ;
400   10.4        11.3744     8.6325 ;
405   10.4        11.3214     8.5761 ;
410   10.4222     11.4496     8.8000 ;
415   10.4222     11.5829     8.7658 ;
420   10.5111     11.588      8.7897 ;
425   10.5111     11.6752     8.9077 ;
430   10.5111     11.735      8.8393 ;
435   10.6667     11.9436     9.0530 ;
440   10.7556     12.0291     9.0427 ;
445   10.7111     12.0667     9.2068 ;
450   10.6889     12.0376     9.0735 ;
**;
```

```
* Identify Parameters to be Varied;
SETVARY ki ke kb kmi kmo kdl kdh v0 BT Vm1 Vm2 Km1 Km2;

BEGIN;

COMPILE INSTANT ;
WRITE 1=10, " PRINT STREAM NO. " % ;
WRITE 2=10, "TIME Med Med(obs) Nuc Nuc(obs) Cyt Cyt(obs)" % ;
WRITE 1=11, " PRINT STREAM NO. " % ;
WRITE 2=11, "TIME Med Med(obs) Nuc Nuc(obs) Cyt Cyt(obs)" % ;
#COUNT=0;
**;

COMPILE PRINT;
* Output routine called during the final phase;
#COUNT=#COUNT+1;
DO 10 FOR #2=#COUNT-1 ;
WRITE 1=10, ((E14,6)) TIME, Med, VOBS<0,#2> ,Nuc, VOBS<1,#2>
Cyt, VOBS<2,#2> ;
WRITE 1=11, ((E14,6)) TIME, Med, VOBS<0,#2> ,Nuc, VOBS<1,#2>
Cyt, VOBS<2,#2> ;
LABEL 10 ;
**;

SETNOFIT;
WHENEVER TIME = TOBS % CALL PRINT;
**;

BEGIN;
STOP;
```

High loading cell

```
* Active Pump Parameter Fitting
*========================================================= ;
* Mohammed Isam Atari                                      ;
*========================================================= ;

EXECUTE ;

OPEN 8 "Pstream1In.txt";
OPEN 9 "Pstream2In.txt";
OPEN 10 "fplot.out";
OPEN 11 "fit_plot_data.txt";
**;


* Define the Variables ;
VARIABLE Lm Hm Le He Lc Hc Ln;

* Define the Parameters ;
PARAMETER
ki ke kb kmi kmo kom kcm koc kcc kdl kdh
v0 v1 v2 v3
D
Vm1 Km1 Vm2 Km2
```

```
alpha 2.4375e4
BT Med Cyt Nuc
;

* Open file, read in and assign parameters ;
COMPILE INSTANT;
OPEN 7 "pars.txt";
READ 7 kom ;
READ 7 kcm ;
READ 7 koc ;
READ 7 kcc ;
READ 7 ki ;
READ 7 ke ;
READ 7 kb ;
READ 7 kmi ;
READ 7 kmo ;
READ 7 kdl ;
READ 7 kdh ;
READ 7 v0 ;
READ 7 v2 ;
READ 7 v3 ;
READ 7 D ;
READ 7 Vm1 ;
READ 7 Km1 ;
READ 7 Vm2 ;
READ 7 Km2 ;
READ 7 BT ;
**;

COMPILE GENERAL ;
Med = Le + He ;
Nuc = Ln ;
Cyt = Lc + Hc ;
**;


INTEGER #COUNT ;

* Assign Intital Conditions for Variables ;
COMPILE INITIAL;
Lm = (1+v0)*D ;
Hm = 0 ;
Le = 0 ;
He = 0 ;
Lc = 0 ;
Hc = 0 ;
Ln = 0 ;
v1 = alpha*v0/(1+v0) ;
**;

COMPILE EQUATIONS;
'Lm = -(kom + kmi)*Lm + kcm*Hm + kmo*v0*Le ;
'Hm = kom*Lm -(kcm + kmi)*Hm + kmo*v0*He ;
'Le = (kmi*Lm)/v0 -(kmo + kom + ki)*Le + kcm*He +
(ke*Lc)/v1 ;
'He = (kmi*Hm)/v0 + kom*Le -(kcm + kmo)*He +
(Vm1*v3*Hc)/((v1)*(v3*Km1 + Hc)) ;
'Lc = ki*v1*Le -(ke + koc)*Lc + kcc*Hc + kdl*v2*Ln - kb*(BT - Ln)*Lc
- (Vm2*Lc)/(Km2 + Lc) ;
'Hc = koc*Lc - kcc*Hc + kdh*v2*Ln - (v3*Vm1*Hc)/(v3*Km1 + Hc)
+ (Vm2*Lc)/(Km2 + Lc) ;
```

```
'Ln = kb*Lc*(BT - Ln)/v2 - (kdl + kdh)*Ln;
**;


SETPSTREAM 1 8;
**;
SETPSTREAM 2 9;
**;

DATA 0.05;
TIME Med Nuc Cyt ;
RANGE  10.8 16.4 12.5 ;
 5       3.4     0.1     0.5 ;
 10      5.9     0.2     0.9 ;
 15      7       0.4     1.3 ;
 20      7.5     0.6     1.3 ;
 25      7.9     1       1.5 ;
 30      8.3     1.3     1.8 ;
 35      8.6     1.6     2 ;
 40      8.7     2       2.2 ;
 45      8.9     2.3     2.7 ;
 50      9       2.6     2.8 ;
 55      9       3       2.7 ;
 60      9       3.2     3.1 ;
 65      9.2     3.5     3.5 ;
 70      9.2     3.8     3.6 ;
 75      9.2     4.2     4 ;
 80      9.2     4.5     4.2 ;
 85      9.3     4.8     4.5 ;
 90      9.3     5       4.5 ;
 95      9.4     5.2     4.4 ;
 100     9.4     5.4     4.9 ;
 105     9.4     5.7     5 ;
 110     9.4     6       5.4 ;
 115     9.4     6.2     5.3 ;
 120     9.4     6.4     5.4 ;
 125     9.4     6.7     5.3 ;
 130     9.4     6.9     5.2 ;
 135     9.5     7.2     5.9 ;
 140     9.6     7.3     5.7 ;
 145     9.5     7.5     6 ;
 150     9.6     7.9     6.2 ;
 155     9.6     8       5.8 ;
 160     9.6     8.2     6.4 ;
 165     9.7     8.4     6.6 ;
 170     9.7     8.5     6.8 ;
 175     9.8     8.9     6.8 ;
 180     9.8     9       7.3 ;
 185     9.8     9.2     7 ;
 190     9.9     9.4     7.3 ;
 195     9.9     9.6     7.5 ;
 200     9.8     9.6     7.2 ;
 205     9.7     9.7     7.5 ;
 210     9.8     10.1    8.2 ;
 215     9.8     10.4    7.8 ;
 220     9.8     10.3    8.1 ;
 225     9.8     10.8    8.1 ;
 230     9.8     10.7    8.3 ;
 235     9.8     11.2    8.7 ;
 240     9.9     11.4    8.4 ;
 245     9.8     11.2    8.7 ;
```

```
 250     9.9     11.3    8.8 ;
 255     9.9     11.5    8.9 ;
 260     9.9     11.8    8.9 ;
 265     10      12      9.4 ;
 270     10      12      9.2 ;
 275     10      12.4    9.6 ;
 280     10      12.3    9.2 ;
 285     10      12.5    9.4 ;
 290     10.1    12.9    9.3 ;
 295     10.2    13      9.8 ;
 300     10.1    12.9    10.1 ;
 305     10.2    13.3    10.3 ;
 310     10.2    13.1    9.7 ;
 315     10.3    13.6    10.4 ;
 320     10.2    13.6    10.2 ;
 325     10.2    13.6    10.3 ;
 330     10.2    13.7    10.4 ;
 335     10.2    14      10 ;
 340     10.3    14.2    10.8 ;
 345     10.3    14.2    10.6 ;
 350     10.3    14.4    10.9 ;
 355     10.3    14.6    10.6 ;
 360     10.3    14.6    10.6 ;
 365     10.4    14.9    11.1 ;
 370     10.4    14.8    10.9 ;
 375     10.4    14.9    10.8 ;
 380     10.4    14.8    11.1 ;
 385     10.4    15.1    11.5 ;
 390     10.4    15.1    11.2 ;
 395     10.4    15.1    11.7 ;
 400     10.4    15.5    11.5 ;
 405     10.4    15.3    11.5 ;
 410     10.4    15.4    12.4 ;
 415     10.4    15.8    12.1 ;
 420     10.5    15.5    11.6 ;
 425     10.5    15.6    12 ;
 430     10.5    16      12 ;
 435     10.7    16.2    12.4 ;
 440     10.8    16.4    12.5 ;
 445     10.7    16.4    12.4 ;
 450     10.7    16.2    11.6 ;
 **;



* Identify Parameters to be Varied;
SETVARY ki ke kb kmi kmo kdl kdh v0 BT Vm1 Vm2 Km1 Km2;

BEGIN;

COMPILE INSTANT ;
WRITE 1=10, " PRINT STREAM NO. " % ;
WRITE 2=10, "TIME Med Med(obs) Nuc Nuc(obs) Cyt Cyt(obs)" % ;
WRITE 1=11, " PRINT STREAM NO. " % ;
WRITE 2=11, "TIME Med Med(obs) Nuc Nuc(obs) Cyt Cyt(obs)" % ;
#COUNT=0;
**;

COMPILE PRINT;
* Output routine called during the final phase;
#COUNT=#COUNT+1;
```

```
DO 10 FOR #2=#COUNT-1 ;
WRITE 1=10, ((E14,6)) TIME, Med, VOBS<0,#2> ,Nuc, VOBS<1,#2>
Cyt, VOBS<2,#2> ;
WRITE 1=11, ((E14,6)) TIME, Med, VOBS<0,#2> ,Nuc, VOBS<1,#2>
Cyt, VOBS<2,#2> ;
LABEL 10 ;
**;

SETNOFIT;
WHENEVER TIME = TOBS % CALL PRINT;
**;

BEGIN;
STOP;
```

Low loading cell

```
* Active Pump Parameter Fitting
*========================================================= ;
* Mohammed Isam Atari                                      ;
*========================================================= ;

EXECUTE ;

OPEN 8 "Pstream1In.txt";
OPEN 9 "Pstream2In.txt";
OPEN 10 "fplot.out";
OPEN 11 "fit_plot_data.txt";
**;


* Define the Variables ;
VARIABLE Lm Hm Le He Lc Hc Ln;

* Define the Parameters ;
PARAMETER
ki ke kb kmi kmo kom kcm koc kcc kdl kdh
v0 v1 v2 v3
D
Vm1 Km1 Vm2 Km2
alpha 2.4375e4
BT Med Cyt Nuc
;

* Open file, read in and assign parameters ;
COMPILE INSTANT;
OPEN 7 "pars.txt";
READ 7 kom ;
READ 7 kcm ;
READ 7 koc ;
READ 7 kcc ;
READ 7 ki ;
READ 7 ke ;
READ 7 kb ;
READ 7 kmi ;
READ 7 kmo ;
READ 7 kdl ;
```

```
READ 7 kdh ;
READ 7 v0 ;
READ 7 v2 ;
READ 7 v3 ;
READ 7 D ;
READ 7 Vm1 ;
READ 7 Km1 ;
READ 7 Vm2 ;
READ 7 Km2 ;
READ 7 BT ;
**;

COMPILE GENERAL ;
Med = Le + He ;
Nuc = Ln ;
Cyt = Lc + Hc ;
**;


INTEGER #COUNT ;

* Assign Intital Conditions for Variables ;
COMPILE INITIAL;
Lm = (1+v0)*D ;
Hm = 0 ;
Le = 0 ;
He = 0 ;
Lc = 0 ;
Hc = 0 ;
Ln = 0 ;
v1 = alpha*v0/(1+v0) ;
**;

COMPILE EQUATIONS;
'Lm = -(kom + kmi)*Lm + kcm*Hm + kmo*v0*Le ;
'Hm = kom*Lm -(kcm + kmi)*Hm + kmo*v0*He ;
'Le = (kmi*Lm)/v0 -(kmo + kom + ki)*Le + kcm*He +
(ke*Lc)/v1 ;
'He = (kmi*Hm)/v0 + kom*Le -(kcm + kmo)*He +
(Vm1*v3*Hc)/((v1)*(v3*Km1 + Hc)) ;
'Lc = ki*v1*Le -(ke + koc)*Lc + kcc*Hc + kdl*v2*Ln - kb*(BT - Ln)*Lc
- (Vm2*Lc)/(Km2 + Lc) ;
'Hc = koc*Lc - kcc*Hc + kdh*v2*Ln - (v3*Vm1*Hc)/(v3*Km1 + Hc)
+ (Vm2*Lc)/(Km2 + Lc) ;
'Ln = kb*Lc*(BT - Ln)/v2 - (kdl + kdh)*Ln;
**;


SETPSTREAM 1 8;
**;
SETPSTREAM 2 9;
**;

DATA 0.05;
TIME Med Nuc Cyt ;
RANGE  10.8 7.4 6.2 ;
 5       3.4     1e-36   0.1;
 10      5.9     0.1     0.1;
 15      7       0.2     0.2;
 20      7.5     0.3     0.4;
 25      7.9     0.5     0.6;
```

```
30      8.3     0.6     0.6;
35      8.6     0.7     0.7;
40      8.7     0.9     0.9;
45      8.9     1.1     0.9;
50      9       1.2     1.1;
55      9       1.4     1.1;
60      9       1.5     1.3;
65      9.2     1.7     1.6;
70      9.2     1.8     1.7;
75      9.2     2.0     1.6;
80      9.2     2.1     1.7;
85      9.3     2.3     1.9;
90      9.3     2.5     1.9;
95      9.4     2.5     2.1;
100     9.4     2.6     2.2;
105     9.4     2.8     2.3;
110     9.4     2.9     2.4;
115     9.4     3.0     2.4;
120     9.4     3.1     2.4;
125     9.4     3.2     2.5;
130     9.4     3.3     2.7;
135     9.5     3.4     2.8;
140     9.6     3.5     2.9;
145     9.5     3.7     2.7;
150     9.6     3.7     3.0;
155     9.6     3.7     3.0;
160     9.6     3.9     3.1;
165     9.7     4.1     3.2;
170     9.7     4.1     3.2;
175     9.8     4.2     3.2;
180     9.8     4.3     3.4;
185     9.8     4.4     3.5;
190     9.9     4.5     3.4;
195     9.9     4.6     3.6;
200     9.8     4.6     3.9;
205     9.7     4.6     3.6;
210     9.8     4.7     4.0;
215     9.8     4.8     3.7;
220     9.8     5.0     3.7;
225     9.8     5.0     3.9;
230     9.8     5.1     4.0;
235     9.8     5.1     4.0;
240     9.9     5.0     4.3;
245     9.8     5.2     4.4;
250     9.9     5.3     4.3;
255     9.9     5.3     4.3;
260     9.9     5.4     4.5;
265     10      5.5     4.4;
270     10      5.6     4.4;
275     10      5.8     4.7;
280     10      5.7     4.5;
285     10      5.8     4.6;
290     10.1    5.8     4.5;
295     10.2    6.0     4.8;
300     10.1    6.0     4.8;
305     10.2    6.0     4.9;
310     10.2    6.1     4.9;
315     10.3    6.2     4.9;
320     10.2    6.2     5.1;
325     10.2    6.2     5.0;
330     10.2    6.2     5.1;
```

```
 335     10.2    6.4     5.2;
 340     10.3    6.4     4.9;
 345     10.3    6.3     5.3;
 350     10.3    6.5     5.2;
 355     10.3    6.6     5.2;
 360     10.3    6.5     5.2;
 365     10.4    6.6     5.3;
 370     10.4    6.7     5.4;
 375     10.4    6.7     5.3;
 380     10.4    6.7     5.6;
 385     10.4    6.8     5.2;
 390     10.4    6.9     5.4;
 395     10.4    6.9     5.7;
 400     10.4    6.9     5.3;
 405     10.4    6.8     5.5;
 410     10.4    7.0     5.6;
 415     10.4    7.1     5.7;
 420     10.5    6.9     5.6;
 425     10.5    7.0     5.9;
 430     10.5    7.2     5.7;
 435     10.7    7.3     6.0;
 440     10.8    7.4     6.2;
 445     10.7    7.4     5.9;
 450     10.7    7.4     5.9;
 **;



* Identify Parameters to be Varied;
SETVARY ki ke kb kmi kmo kdl kdh v0 BT Vm1 Vm2 Km1 Km2;

BEGIN;

COMPILE INSTANT ;
WRITE 1=10, " PRINT STREAM NO. " % ;
WRITE 2=10, "TIME Med Med(obs) Nuc Nuc(obs) Cyt Cyt(obs)" % ;
WRITE 1=11, " PRINT STREAM NO. " % ;
WRITE 2=11, "TIME Med Med(obs) Nuc Nuc(obs) Cyt Cyt(obs)" % ;
#COUNT=0;
**;

COMPILE PRINT;
* Output routine called during the final phase;
#COUNT=#COUNT+1;
DO 10 FOR #2=#COUNT-1 ;
WRITE 1=10, ((E14,6)) TIME, Med, VOBS<0,#2> ,Nuc, VOBS<1,#2>
Cyt, VOBS<2,#2> ;
WRITE 1=11, ((E14,6)) TIME, Med, VOBS<0,#2> ,Nuc, VOBS<1,#2>
Cyt, VOBS<2,#2> ;
LABEL 10 ;
**;

SETNOFIT;
WHENEVER TIME = TOBS % CALL PRINT;
**;

BEGIN;
STOP;
```

Simultaneous fitting

```
* Active Pump Parameter Fitting
*========================================================= ;
* Mohammed Isam Atari                                      ;
*========================================================= ;

EXECUTE ;
OPEN 8 "fplot_c.out";
OPEN 9 "fplot_r.out";
OPEN 10 "fplot_p1.out";
OPEN 11 "fplot_p2.out";
**;



* Define the Variables ;
VARIABLE Lmh Hmh Leh Heh Lch Hch Lnh
Lml Hml Lel Hel Lcl Hcl Lnl ;

* Define the Parameters ;
PARAMETER
ki ke kb kmi kmo kom kcm koc kcc kdl kdh
v0 v1 v2 v3
D
Vm1h Km1h Vm2h Km2h
Vm1l Km1l Vm2l Km2l
alpha 2.4375e4
BT Medh Cyth Nuch
Medl Cytl Nucl
;

* Open file, read in and assign parameters ;
COMPILE INSTANT;
OPEN 7 "pars.txt";
READ 7 kom ;
READ 7 kcm ;
READ 7 koc ;
READ 7 kcc ;
READ 7 ki ;
READ 7 ke ;
READ 7 kb ;
READ 7 kmi ;
READ 7 kmo ;
READ 7 kdl ;
READ 7 kdh ;
READ 7 v0 ;
READ 7 v2 ;
READ 7 v3 ;
READ 7 D ;
READ 7 Vm1h ;
READ 7 Km1h ;
READ 7 Vm2h ;
READ 7 Km2h ;
READ 7 Vm1l ;
READ 7 Km1l ;
READ 7 Vm2l ;
READ 7 Km2l ;
```

```
READ 7 BT ;
**;

COMPILE GENERAL ;
Medh = Leh + Heh ;
Nuch = Lnh ;
Cyth = Lch + Hch ;
Medl = Lel + Hel ;
Nucl = Lnl ;
Cytl = Lcl + Hcl ;
**;


INTEGER #COUNT ;

* Assign Intital Conditions for Variables ;
COMPILE INITIAL;
Lmh = (1+v0)*D ;
Hmh = 0 ;
Leh = 0 ;
Heh = 0 ;
Lch = 0 ;
Hch = 0 ;
Lnh = 0 ;
Lml = (1+v0)*D ;
Hml = 0 ;
Lel = 0 ;
Hel = 0 ;
Lcl = 0 ;
Hcl = 0 ;
Lnl = 0 ;
v1 = alpha*v0/(1+v0) ;
**;

COMPILE EQUATIONS;
'Lmh = -(kom + kmi)*Lmh + kcm*Hmh + kmo*v0*Leh ;
'Hmh = kom*Lmh -(kcm + kmi)*Hmh + kmo*v0*Heh ;
'Leh = (kmi*Lmh)/v0 -(kmo + kom + ki)*Leh + kcm*Heh +
(ke*Lch)/v1 ;
'Heh = (kmi*Hmh)/v0 + kom*Leh -(kcm + kmo)*Heh +
(Vm1h*v3*Hch)/((v1)*(v3*Km1h + Hch)) ;
'Lch = ki*v1*Leh -(ke + koc)*Lch + kcc*Hch + kdl*v2*Lnh
- kb*(BT - Lnh)*Lch - (Vm2h*Lch)/(Km2h + Lch) ;
'Hch = koc*Lch - kcc*Hch + kdh*v2*Lnh - (v3*Vm1h*Hch)/(v3*Km1h + Hch)
+ (Vm2h*Lch)/(Km2h + Lch) ;
'Lnh = kb*Lch*(BT - Lnh)/v2 - (kdl + kdh)*Lnh;
'Lml = -(kom + kmi)*Lml + kcm*Hml + kmo*v0*Lel ;
'Hml = kom*Lml -(kcm + kmi)*Hml + kmo*v0*Hel ;
'Lel = (kmi*Lml)/v0 -(kmo + kom + ki)*Lel + kcm*Hel +
(ke*Lcl)/v1 ;
'Hel = (kmi*Hml)/v0 + kom*Lel -(kcm + kmo)*Hel +
(Vm1l*v3*Hcl)/((v1)*(v3*Km1l + Hcl)) ;
'Lcl = ki*v1*Lel -(ke + koc)*Lcl + kcc*Hcl + kdl*v2*Lnl
- kb*(BT - Lnl)*Lcl - (Vm2l*Lcl)/(Km2l + Lcl) ;
'Hcl = koc*Lcl - kcc*Hcl + kdh*v2*Lnl - (v3*Vm1l*Hcl)/(v3*Km1l + Hcl)
+ (Vm2l*Lcl)/(Km2l + Lcl) ;
'Lnl = kb*Lcl*(BT - Lnl)/v2 - (kdl + kdh)*Lnl ;
**;


SETPSTREAM 1 8;
```

```
**;
SETPSTREAM 2 9;
**;

DATA 0.05;
TIME Medh Nuch Cyth Medl Nucl Cytl ;
RANGE  10.76 16.4 12.51 10.76 7.44 6.18;
5     3.40  0.07  0.51  3.40  0.04  0.11  ;
10    5.89  0.20  0.91  5.89  0.11  0.09  ;
15    6.98  0.44  1.29  6.98  0.20  0.22  ;
20    7.53  0.62  1.29  7.53  0.33  0.38  ;
25    7.91  0.96  1.49  7.91  0.47  0.56  ;
30    8.29  1.31  1.78  8.29  0.62  0.60  ;
35    8.58  1.62  2.00  8.58  0.73  0.73  ;
40    8.71  1.96  2.22  8.71  0.91  0.87  ;
45    8.87  2.31  2.69  8.87  1.09  0.93  ;
50    9.00  2.58  2.78  9.00  1.22  1.07  ;
55    9.00  2.96  2.71  9.00  1.42  1.11  ;
60    9.02  3.18  3.13  9.02  1.53  1.29  ;
65    9.18  3.53  3.51  9.18  1.73  1.56  ;
70    9.22  3.80  3.64  9.22  1.82  1.67  ;
75    9.20  4.22  4.04  9.20  2.00  1.64  ;
80    9.22  4.51  4.16  9.22  2.13  1.69  ;
85    9.29  4.76  4.51  9.29  2.31  1.91  ;
90    9.31  4.98  4.49  9.31  2.47  1.93  ;
95    9.36  5.20  4.42  9.36  2.49  2.07  ;
100   9.44  5.42  4.89  9.44  2.64  2.20  ;
105   9.38  5.69  4.96  9.38  2.78  2.29  ;
110   9.40  5.98  5.40  9.40  2.87  2.38  ;
115   9.44  6.18  5.31  9.44  3.02  2.42  ;
120   9.42  6.40  5.42  9.42  3.13  2.42  ;
125   9.40  6.71  5.27  9.40  3.20  2.51  ;
130   9.44  6.93  5.24  9.44  3.33  2.73  ;
135   9.47  7.16  5.93  9.47  3.38  2.76  ;
140   9.56  7.27  5.71  9.56  3.51  2.91  ;
145   9.53  7.47  6.00  9.53  3.69  2.67  ;
150   9.60  7.89  6.16  9.60  3.71  3.02  ;
155   9.56  7.98  5.78  9.56  3.73  3.04  ;
160   9.62  8.18  6.40  9.62  3.87  3.13  ;
165   9.73  8.38  6.60  9.73  4.07  3.18  ;
170   9.73  8.53  6.76  9.73  4.11  3.22  ;
175   9.80  8.91  6.84  9.80  4.18  3.20  ;
180   9.78  8.96  7.31  9.78  4.27  3.40  ;
185   9.78  9.16  7.04  9.78  4.38  3.47  ;
190   9.87  9.36  7.33  9.87  4.47  3.44  ;
195   9.87  9.62  7.49  9.87  4.62  3.62  ;
200   9.78  9.60  7.16  9.78  4.60  3.93  ;
205   9.73  9.67  7.49  9.73  4.62  3.58  ;
210   9.76  10.07 8.22 9.76 4.73 4.00 ;
215   9.76  10.38 7.82 9.76 4.76 3.67 ;
220   9.84  10.33 8.09  9.84  5.00  3.73  ;
225   9.82  10.80 8.11  9.82  5.02  3.91  ;
230   9.76  10.73 8.31  9.76  5.07  3.96  ;
235   9.80  11.18 8.69  9.80  5.11  4.04  ;
240   9.89  11.38 8.36  9.89  5.02  4.33  ;
245   9.84  11.18 8.71  9.84  5.22  4.42  ;
250   9.91  11.33 8.80  9.91  5.31  4.33  ;
255   9.89  11.53 8.93  9.89  5.33  4.31  ;
260   9.93  11.76 8.91  9.93  5.44  4.49  ;
265   9.96  12.00 9.44  9.96  5.53  4.42  ;
270   10.00 12.00 9.20  10.00 5.58  4.40  ;
```

```
275   10.00 12.40 9.58  10.00 5.76  4.71  ;
280   10.02 12.33 9.22  10.02 5.71  4.53  ;
285   10.02 12.53 9.44  10.02 5.78  4.58  ;
290   10.09 12.87 9.27  10.09 5.82  4.49  ;
295   10.16 13.02 9.78  10.16 5.98  4.84  ;
300   10.13 12.93 10.13 10.13 5.96  4.78  ;
305   10.18 13.27 10.29 10.18 6.04  4.89  ;
310   10.18 13.09 9.69  10.18 6.13  4.89  ;
315   10.29 13.56 10.44 10.29 6.16  4.89  ;
320   10.22 13.56 10.16 10.22 6.22  5.13  ;
325   10.22 13.64 10.29 10.22 6.22  5.02  ;
330   10.24 13.67 10.44 10.24 6.20  5.09  ;
335   10.22 13.96 10.02 10.22 6.36  5.16  ;
340   10.29 14.24 10.78 10.29 6.38  4.87  ;
345   10.29 14.18 10.62 10.29 6.33  5.31  ;
350   10.33 14.40 10.93 10.33 6.49  5.20  ;
355   10.31 14.56 10.56 10.31 6.62  5.24  ;
360   10.33 14.56 10.56 10.33 6.53  5.24  ;
365   10.38 14.91 11.13 10.38 6.58  5.27  ;
370   10.42 14.76 10.93 10.42 6.73  5.42  ;
375   10.42 14.87 10.78 10.42 6.71  5.33  ;
380   10.36 14.84 11.09 10.36 6.73  5.56  ;
385   10.42 15.07 11.51 10.42 6.76  5.24  ;
390   10.40 15.11 11.22 10.40 6.93  5.36  ;
395   10.38 15.09 11.71 10.38 6.87  5.71  ;
400   10.40 15.51 11.51 10.40 6.93  5.29  ;
405   10.40 15.31 11.47 10.40 6.78  5.47  ;
410   10.42 15.42 12.38 10.42 7.00  5.62  ;
415   10.42 15.80 12.11 10.42 7.11  5.71  ;
420   10.51 15.53 11.60 10.51 6.93  5.56  ;
425   10.51 15.60 12.04 10.51 7.04  5.93  ;
430   10.51 16.02 11.96 10.51 7.16  5.71  ;
435   10.67 16.18 12.44 10.67 7.27  6.00  ;
440   10.76 16.40 12.51 10.76 7.36  6.18  ;
445   10.71 16.40 12.42 10.71 7.36  5.89  ;
450   10.69 16.16 11.64 10.69 7.44  5.87  ;
**;



* Identify Parameters to be Varied;
SETVARY ki ke kb kmi kmo kdl kdh v0 BT Vm1h
Vm2h Km1h Km2h Vm1l Vm2l Km1l Km2l;

BEGIN;

COMPILE INSTANT ;
WRITE 1=10, " PRINT STREAM NO. " % ;
WRITE 1, "TIME Medh Medh(obs) Nuch Nuch(obs) Cyth Cyth(obs) " % ;
WRITE 1=11, " PRINT STREAM NO. " % ;
WRITE 1, "TIME Medl Medl(obs) Nucl Nucl(obs) Cytl Cytl(obs)" % ;
#COUNT=0;
**;

COMPILE PRINT;
* Output routine called during the final phase;
#COUNT=#COUNT+1;
DO 10 FOR #2=#COUNT-1 ;
WRITE 1=10, ((E14,6)) TIME, Medh, VOBS<0,#2>, Nuch, VOBS<1,#2>,
Cyth, VOBS<2,#2> ;
WRITE 1=11, ((E14,6)) TIME, Medl, VOBS<3,#2>, Nucl, VOBS<4,#2>,
```

```
Cytl, VOBS<5,#2> ;
LABEL 10 ;
**;

SETNOFIT;
WHENEVER TIME = TOBS % CALL PRINT;
**;

BEGIN;
STOP;
```

Parameter estimation of the extended cell cycle model

```
* Extended Cell Cycle Model Parameter Estimation;


EXECUTE ;
OPEN 8 "Pstream1In.txt" ;
OPEN 9 "Pstream2In.txt" ;
OPEN 10 "plot.out";
OPEN 11 "fit_plot_data.txt";
**;


VARIABLE cyclin cdc2cyclin cdc2cyclinyp cdc2cyclinyptp
cdc2cyclintp cdc2 cdc25act
wee1act apcstar apc0 p21 ;


* Define the Parameters ;
PARAMETER
ksynth kdest ka kd factor
kwee1 kcdc25
cdc2tot cdc25tot wee1tot apctot p21tot
ec50wee1 ec50cdc25 ec50apc ec50p21
kcdc25on kcdc25off kapcon kapcoff kwee1on kwee1off
kcak kpp2c C2 kp21on kp21off
alpha beta CycBact C1
;

INTEGER #nwee1 #ncdc25 #napc #np21 #COUNT ;

* Open file, read in and assign parameters ;
COMPILE INSTANT;
OPEN 7 "pars1.txt";
READ 7 ksynth ;
READ 7 kdest ;
READ 7 ka ;
READ 7 kd ;
READ 7 factor ;
READ 7 kwee1 ;
READ 7 kcdc25 ;
READ 7 cdc2tot ;
READ 7 cdc25tot ;
READ 7 wee1tot ;
READ 7 apctot ;
```

```
READ 7 #nwee1 ;
READ 7 #ncdc25 ;
READ 7 #napc ;
READ 7 ec50wee1 ;
READ 7 ec50cdc25 ;
READ 7 ec50apc ;
READ 7 kcdc25on ;
READ 7 kcdc25off ;
READ 7 kapcon ;
READ 7 kapcoff ;
READ 7 kwee1on ;
READ 7 kwee1off ;
READ 7 kcak ;
READ 7 kpp2c ;
READ 7 C2 ;
READ 7 kp21on ;
READ 7 kp21off ;
READ 7 #np21 ;
READ 7 p21tot ;
READ 7 ec50p21 ;
READ 7 alpha ;
READ 7 beta ;
READ 7 C1 ;
**;

COMPILE GENERAL ;
CycBact = alpha*cdc2cyclintp + beta ;
**;

* Assign intital conditions for variables ;
COMPILE INITIAL;
cyclin = C2*0.141396 ;
cdc2cyclin = C2*0.992654 ;
cdc2cyclinyp = C2*1.4763 ;
cdc2cyclinyptp = C2*11.41023 ;
cdc2cyclintp = C2*1.64121 ;
cdc2 = C2*84.479606 ;
cdc25act = C2*1.54009 ;
wee1act = C2*15 ;
apc0 = C2*46.1461 ;
apcstar = C2*3.8539 ;
p21 = C2*3.8539 ;
**;



COMPILE EQUATIONS;
'cyclin = C1*ksynth - C1*kdest*(apcstar + p21)*cyclin -
C1*ka*cdc2*cyclin + C1*kd*cdc2cyclin ;
'cdc2cyclin = C1*ka*cdc2*cyclin - C1*kd*cdc2cyclin -
C1*kdest*(apcstar + p21)*cdc2cyclin -
C1*kwee1*wee1act*cdc2cyclin -
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclin +
C1*kcdc25*cdc25act*cdc2cyclinyp +
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyp ;
'cdc2cyclinyp = C1*kwee1*wee1act*cdc2cyclin +
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclin -
C1*kcdc25*cdc25act*cdc2cyclinyp -
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyp -
C1*kcak*cdc2cyclinyp + C1*kpp2c*cdc2cyclinyptp -
C1*kdest*(apcstar + p21)*cdc2cyclinyp ;
```

```
'cdc2cyclinyptp = C1*kcak*cdc2cyclinyp -
C1*kpp2c*cdc2cyclinyptp -
C1*kcdc25*cdc25act*cdc2cyclinyptp -
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyptp +
C1*kwee1*wee1act*cdc2cyclintp +
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclintp -
C1*kdest*(apcstar + p21)*cdc2cyclinyptp ;
'cdc2cyclintp = C1*kcdc25*cdc25act*cdc2cyclinyptp +
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyptp -
C1*kwee1*wee1act*cdc2cyclintp -
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclintp -
C1*kdest*(apcstar + p21)*cdc2cyclintp ;
'cdc2 = C1*kd*cdc2cyclin - C1*ka*cdc2*cyclin +
C1*kdest*(apcstar + p21)*(cdc2cyclin + cdc2cyclinyp +
cdc2cyclinyptp + cdc2cyclintp) ;
'cdc25act = -C1*kcdc25off*cdc25act +
C1*kcdc25on*(C2*cdc25tot - cdc25act)*((cdc2cyclintp**#ncdc25)/
(C2*ec50cdc25**#ncdc25 + cdc2cyclintp**#ncdc25)) ;
'wee1act = C1*kwee1on*(C2*wee1tot - wee1act) -
wee1act*C1*kwee1off*((cdc2cyclintp**#nwee1)/
(C2*ec50wee1**#nwee1 + cdc2cyclintp**#nwee1)) ;
'apcstar = - C1*kapcoff*apcstar +
C1*kapcon*(C2*apctot - apcstar)*((cdc2cyclintp**#napc)/
(C2*ec50apc**#napc + cdc2cyclintp**#napc)) ;
'apc0 = C1*kapcoff*apcstar -
C1*kapcon*(C2*apctot - apcstar)*((cdc2cyclintp**#napc)/
(C2*ec50apc**#napc + cdc2cyclintp**#napc)) ;
'p21 = - C1*kp21off*p21 +
C1*kp21on*(C2*p21tot - p21)*((cdc2cyclintp**#np21)/
(C2*ec50p21**#np21 + cdc2cyclintp**#np21)) ;
**;

SETPSTREAM 1 8;
**;
SETPSTREAM 2 9;
**;


* DATA relative error estimate default 0.01;
DATA 0.1 ;
TIME CycBact ;
RANGE 345.05 ;
20    119.01      ;
40    97.085      ;
60    92.51 ;
80    85.175      ;
100   83.575      ;
120   80.4  ;
140   86.96 ;
160   86.86 ;
180   91.86 ;
200   92.44 ;
220   85.19 ;
240   88.53 ;
260   94.465      ;
280   114.77      ;
300   117.07      ;
320   125.82      ;
340   126.38      ;
360   142.12      ;
380   131.65      ;
```

```
400   147.84      ;
420   115.62      ;
440   118.81      ;
460   123.65      ;
480   130.54      ;
500   141.22      ;
520   134.5 ;
540   131.95      ;
560   145.21      ;
580   153.42      ;
600   148.16      ;
620   140.33      ;
640   140.84      ;
660   144.17      ;
680   155.25      ;
700   158.17      ;
720   167.97      ;
740   170.97      ;
760   170.87      ;
780   179.77      ;
800   174.8 ;
820   159.13      ;
840   160.6 ;
860   145.67      ;
880   143.03      ;
900   167.66      ;
920   166.83      ;
940   200.41      ;
960   197.27      ;
980   198.44      ;
1000  213.57      ;
1020  199.8 ;
1040  232.64      ;
1060  237.55      ;
1080  224.57      ;
1100  226.65      ;
1120  218.78      ;
1140  227.57      ;
1160  215.26      ;
1180  225.05      ;
1200  243.9 ;
1220  215.38      ;
1240  222.27      ;
1260  246.39      ;
1280  261.6 ;
1300  345.05      ;
**;

* Identify parameters to be varied, not more than 10;
SETVARY alpha C2 beta kp21off p21tot ;

BEGIN;

COMPILE INSTANT ;
WRITE 1=10, "PRINT STREAM NO." % ;
WRITE 1, "TIME CycBact CycBact(obs)" % ;
#COUNT=0;
**;

COMPILE PRINT;
* Output routine called during the final phase;
```

```
#COUNT=#COUNT+1;
DO 10 FOR #2=#COUNT-1 ;
WRITE 1=10, ((E14,6)) TIME, CycBact, VOBS<0,#2> ;
WRITE 1=11, ((E14,6)) TIME, CycBact, VOBS<0,#2> ;
LABEL 10 ;
**;

SETNOFIT;
WHENEVER TIME = TOBS % CALL PRINT;
**;

BEGIN;
STOP;
```

Parameter estimation of the coupled drug kinetics/cell cycle model

Parameter estimation before the *washout* event

```
*========================================================= ;
* Mohammed Isam Atari                                      ;
*========================================================= ;

EXECUTE ;
OPEN 10 "cell.out";
OPEN 11 "cell.txt";
OPEN 12 "medium.out";
OPEN 13 "medium.txt";
OPEN 14 "H2AX.out";
OPEN 15 "H2AX.txt";
**;

SETHMAX 0.1;

* Define the Variables ;
VARIABLE Lm Hm Le He Lc Hc Ln H2AX;

* Define the Parameters ;
PARAMETER
ki ke kb kmi kmo kom kcm koc kcc kdl kdh
v0 v1 v2 v3
Vm1 Km1 Vm2 Km2
alpha 2.4375e4
BT Med Cyt Nuc
ktpt kdam
D
;

* Open file, read in and assign parameters ;
COMPILE INSTANT;
OPEN 7 "pars1wash.txt";
READ 7 kom ;
```

```
READ 7 kcm ;
READ 7 koc ;
READ 7 kcc ;
READ 7 ki ;
READ 7 ke ;
READ 7 kb ;
READ 7 kmi ;
READ 7 kmo ;
READ 7 kdl ;
READ 7 kdh ;
READ 7 v0 ;
READ 7 v2 ;
READ 7 v3 ;
READ 7 D ;
READ 7 Vm1 ;
READ 7 Km1 ;
READ 7 Vm2 ;
READ 7 Km2 ;
READ 7 BT ;
READ 7 ktpt ;
READ 7 kdam ;
**;


* Assign Intital Conditions for Variables ;
COMPILE INITIAL;
Lm = (1+v0)*D ;
Hm = 0 ;
Le = 0 ;
He = 0 ;
Lc = 0 ;
Hc = 0 ;
Ln = 0 ;
H2AX = 0 ;
v1 = alpha*v0/(1+v0) ;
**;

COMPILE EQUATIONS;
'Lm = -(kom + kmi)*Lm + kcm*Hm + kmo*v0*Le ;
'Hm = kom*Lm -(kcm + kmi)*Hm + kmo*v0*He ;
'Le = (kmi*Lm)/v0 -(kmo + kom + ki)*Le + kcm*He +
(ke*Lc)/v1 ;
'He = (kmi*Hm)/v0 + kom*Le -(kcm + kmo)*He +
(Vm1*v3*Hc)/((v1)*(v3*Km1 + Hc)) ;
'Lc = ki*v1*Le -(ke + koc)*Lc + kcc*Hc + kdl*v2*Ln - kb*(BT - Ln)*Lc
- (Vm2*Lc)/(Km2 + Lc) ;
'Hc = koc*Lc - kcc*Hc + kdh*v2*Ln - (Vm1*v3*Hc)/(v3*Km1 + Hc)
+ (Vm2*Lc)/(Km2 + Lc) ;
'Ln = kb*Lc*(BT - Ln)/v2 - (kdl + kdh)*Ln;
'H2AX = ktpt*Ln - kdam*H2AX ;
**;


COMPILE INSTANT ;
WRITE 1=10, "PRINT STREAM NO." % ;
WRITE 1, "TIME Ln Lc Hc" % ;
WRITE 1=12, "PRINT STREAM NO." % ;
WRITE 1, "TIME Lm Hm Le He" % ;
WRITE 1=14, "PRINT STREAM NO." % ;
WRITE 1, "TIME H2AX" % ;
**;
```

```
COMPILE PRINT;
WRITE 1=10, ((E14,6)) TIME, Ln, Lc, Hc % ;
WRITE 1=11, ((E14,6)) TIME, Ln, Lc, Hc % ;
WRITE 1=12, ((E14,6)) TIME, Lm, Hm, Le, He % ;
WRITE 1=13, ((E14,6)) TIME, Lm, Hm, Le, He % ;
WRITE 1=14, ((E14,6)) TIME, H2AX % ;
WRITE 1=15, ((E14,6)) TIME, H2AX % ;
**;


WHENEVER TIME = 0.0 + 2*1800 % CALL PRINT;
**;
BEGIN;
STOP;
```

Parameter estimation of the coupled model

```
* Coupled drug kinetics/cell cycle response TPT = 1 micromolar;


EXECUTE ;
OPEN 8 "Pstream1In.txt" ;
OPEN 9 "Pstream2In.txt" ;
OPEN 10 "CycBresults.out";
OPEN 11 "CycBresults.txt";
OPEN 12 "Cell.out";
OPEN 13 "Cell.txt";
OPEN 14 "H2AX.out";
OPEN 15 "H2AX.txt";
OPEN 16 "medium.out";
OPEN 17 "medium.txt";
**;



VARIABLE cyclin cdc2cyclin cdc2cyclinyp cdc2cyclinyptp
cdc2cyclintp cdc2 cdc25act
wee1act apcstar apc0 p21
Lm Hm Le He Lc Hc Ln H2AX;


* Define the Parameters ;
PARAMETER
ksynth kdest ka kd factor
kwee1 kcdc25
cdc2tot cdc25tot wee1tot apctot p21tot
ec50wee1 ec50cdc25 ec50apc ec50p21
kcdc25on kcdc25off kapcon kapcoff kwee1on kwee1off
kcak kpp2c C2 kp21on kp21off
alpha beta CycBact C1
ki ke kb kmi kmo kom kcm koc kcc kdl kdh
v0 v1 v2 v3
Vm1 Km1 Vm2 Km2
alphanew 2.4375e4
BT Med Cyt Nuc
ktpt kdam
```

```
;

INTEGER #nwee1 #ncdc25 #napc #np21 #COUNT ;

* Open file, read in and assign parameters ;
COMPILE INSTANT;
OPEN 7 "pars1.txt";
READ 7 ksynth ;
READ 7 kdest ;
READ 7 ka ;
READ 7 kd ;
READ 7 factor ;
READ 7 kwee1 ;
READ 7 kcdc25 ;
READ 7 cdc2tot ;
READ 7 cdc25tot ;
READ 7 wee1tot ;
READ 7 apctot ;
READ 7 #nwee1 ;
READ 7 #ncdc25 ;
READ 7 #napc ;
READ 7 ec50wee1 ;
READ 7 ec50cdc25 ;
READ 7 ec50apc ;
READ 7 kcdc25on ;
READ 7 kcdc25off ;
READ 7 kapcon ;
READ 7 kapcoff ;
READ 7 kwee1on ;
READ 7 kwee1off ;
READ 7 kcak ;
READ 7 kpp2c ;
READ 7 kp21on ;
READ 7 #np21 ;
READ 7 ec50p21 ;
READ 7 C1 ;
READ 7 kom ;
READ 7 kcm ;
READ 7 koc ;
READ 7 kcc ;
READ 7 ki ;
READ 7 ke ;
READ 7 kb ;
READ 7 kmi ;
READ 7 kmo ;
READ 7 kdl ;
READ 7 kdh ;
READ 7 v0 ;
READ 7 v2 ;
READ 7 v3 ;
READ 7 Vm1 ;
READ 7 Km1 ;
READ 7 Vm2 ;
READ 7 Km2 ;
READ 7 BT ;
READ 7 ktpt ;
READ 7 kdam ;
READ 7 alpha ;
READ 7 beta ;
READ 7 kp21off ;
READ 7 C2 ;
```

```
READ 7 p21tot ;
**;

COMPILE GENERAL ;
CycBact = alpha*cdc2cyclintp + beta ;
**;

* Assign intital conditions for variables ;
COMPILE INITIAL;
cyclin = C2*0.141396 ;
cdc2cyclin = C2*0.992654 ;
cdc2cyclinyp = C2*1.4763 ;
cdc2cyclinyptp = C2*11.41023 ;
cdc2cyclintp = C2*1.64121 ;
cdc2 = C2*84.479606 ;
cdc25act = C2*1.54009 ;
wee1act = C2*15 ;
apc0 = C2*46.1461 ;
apcstar = C2*3.8539 ;
p21 = C2*3.8539 ;
Lm = 5.844233E-02 ;
Hm = 4.573399E-02 ;
Le = 6.586645E-02 ;
He = 5.049545E+00 ;
Lc = 1.056672E-02 ;
Hc = 6.97E+03 ;
Ln = 1.971363E-02 ;
H2AX = 2.91 ;
v1 = alphanew*v0/(1+v0) ;
**;



COMPILE EQUATIONS;
'cyclin = C1*ksynth - C1*kdest*(apcstar + p21)*cyclin -
C1*ka*cdc2*cyclin + C1*kd*cdc2cyclin ;
'cdc2cyclin = C1*ka*cdc2*cyclin - C1*kd*cdc2cyclin -
C1*kdest*(apcstar + p21)*cdc2cyclin -
C1*kwee1*wee1act*cdc2cyclin -
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclin +
C1*kcdc25*cdc25act*cdc2cyclinyp +
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyp ;
'cdc2cyclinyp = C1*kwee1*wee1act*cdc2cyclin +
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclin -
C1*kcdc25*cdc25act*cdc2cyclinyp -
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyp -
C1*kcak*cdc2cyclinyp + C1*kpp2c*cdc2cyclinyptp -
C1*kdest*(apcstar + p21)*cdc2cyclinyp ;
'cdc2cyclinyptp = C1*kcak*cdc2cyclinyp -
C1*kpp2c*cdc2cyclinyptp -
C1*kcdc25*cdc25act*cdc2cyclinyptp -
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyptp +
C1*kwee1*wee1act*cdc2cyclintp +
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclintp -
C1*kdest*(apcstar + p21)*cdc2cyclinyptp ;
'cdc2cyclintp = C1*kcdc25*cdc25act*cdc2cyclinyptp +
(C1*kcdc25/factor)*(C2*cdc25tot - cdc25act)*cdc2cyclinyptp -
C1*kwee1*wee1act*cdc2cyclintp -
(C1*kwee1/factor)*(C2*wee1tot - wee1act)*cdc2cyclintp -
C1*kdest*(apcstar + p21)*cdc2cyclintp ;
'cdc2 = C1*kd*cdc2cyclin - C1*ka*cdc2*cyclin +
```

```
C1*kdest*(apcstar + p21)*(cdc2cyclin + cdc2cyclinyp +
cdc2cyclinyptp + cdc2cyclintp) ;
'cdc25act = -C1*kcdc25off*cdc25act +
C1*kcdc25on*(C2*cdc25tot - cdc25act)*((cdc2cyclintp**#ncdc25)/
(C2*ec50cdc25**#ncdc25 + cdc2cyclintp**#ncdc25)) ;
'wee1act = C1*kwee1on*(C2*wee1tot - wee1act) -
wee1act*C1*kwee1off*((cdc2cyclintp**#nwee1)/
(C2*ec50wee1**#nwee1 + cdc2cyclintp**#nwee1)) ;
'apcstar = - C1*kapcoff*apcstar +
C1*kapcon*(C2*apctot - apcstar)*((cdc2cyclintp**#napc)/
(C2*ec50apc**#napc + cdc2cyclintp**#napc)) ;
'apc0 = C1*kapcoff*apcstar -
C1*kapcon*(C2*apctot - apcstar)*((cdc2cyclintp**#napc)/
(C2*ec50apc**#napc + cdc2cyclintp**#napc)) ;
'p21 = - C1*kp21off*p21/(1 + kdam*H2AX) +
C1*kp21on*(C2*p21tot - p21)*((cdc2cyclintp**#np21)/
(C2*ec50p21**#np21 + cdc2cyclintp**#np21)) ;
'Lm = -(kom + kmi)*Lm + kcm*Hm + kmo*v0*Le ;
'Hm = kom*Lm -(kcm + kmi)*Hm + kmo*v0*He ;
'Le = (kmi*Lm)/v0 -(kmo + kom + ki)*Le + kcm*He +
(ke*Lc)/v1 ;
'He = (kmi*Hm)/v0 + kom*Le -(kcm + kmo)*He +
(Vm1*v3*Hc)/((v1)*(v3*Km1 + Hc)) ;
'Lc = ki*v1*Le -(ke + koc)*Lc + kcc*Hc + kdl*v2*Ln - kb*(BT - Ln)*Lc
- (Vm2*Lc)/(Km2 + Lc) ;
'Hc = koc*Lc - kcc*Hc + kdh*v2*Ln - (Vm1*v3*Hc)/(v3*Km1 + Hc)
+ (Vm2*Lc)/(Km2 + Lc) ;
'Ln = kb*Lc*(BT - Ln)/v2 - (kdl + kdh)*Ln;
'H2AX = ktpt*Ln - kdam*H2AX ;
**;

SETPSTREAM 1 8;
**;
SETPSTREAM 2 9;
**;


* DATA relative error estimate default 0.01;
DATA 0.2 ;
TIME CycBact ;
RANGE 66.2 ;
100   45.84 ;
120   48.09 ;
140   47.415      ;
160   40.76 ;
180   44.265      ;
200   40.35 ;
220   40.25 ;
240   41.19 ;
260   41.725      ;
280   39.32 ;
300   39.205      ;
320   37.78 ;
340   38.46 ;
360   38.63 ;
380   42.465      ;
400   44.06 ;
420   44.77 ;
440   38.765      ;
460   36.585      ;
480   36.48 ;
```

```
500   36.41 ;
520   36.28 ;
540   32.495      ;
560   31.43 ;
580   34.665      ;
600   31.445      ;
620   39.95 ;
640   32.705      ;
660   41.46 ;
680   37.78 ;
700   37.13 ;
720   36.665      ;
740   38.85 ;
760   34.08 ;
780   36.3  ;
800   34.885      ;
820   34.675      ;
840   32.47 ;
860   27.225      ;
880   31.61 ;
900   30.385      ;
920   33.58 ;
940   31.985      ;
960   28.155      ;
980   25.32 ;
1000  28.315      ;
1020  32.325      ;
1040  32.315      ;
1060  37.47 ;
1080  33.83 ;
1100  35.965      ;
1120  37.545      ;
1140  36.095      ;
1160  34.005      ;
1180  32.175      ;
1200  30.405      ;
1220  23.295      ;
1240  23.905      ;
1260  25.115      ;
1280  24.065      ;
1300  24.975      ;
1320  22.775      ;
1340  24.035      ;
1360  24.54 ;
1380  26.23 ;
1400  26.99 ;
1420  44.21 ;
1440  43.805      ;
1460  39.7  ;
1480  46.37 ;
1500  45.89 ;
1520  43.035      ;
1540  47.805      ;
1560  48.91 ;
1600  66.2  ;
1620  16.62 ;
1640  15.375      ;
**;

* Identify parameters to be varied, not more than 10;
SETVARY alpha C2 beta kp21off p21tot ;
```

```
BEGIN;

COMPILE INSTANT ;
WRITE 1=10, "PRINT STREAM NO." % ;
WRITE 1, "TIME CycBact CycBact(obs)" % ;
WRITE 1=12, "PRINT STREAM NO." % ;
WRITE 1, "TIME Ln Lc Hc" % ;
WRITE 1=14, "PRINT STREAM NO." % ;
WRITE 1, "TIME H2AX" % ;
WRITE 1=16, "PRINT STREAM NO." % ;
WRITE 1, "TIME Lm Hm Le He" % ;
#COUNT=0;
**;

COMPILE PRINT;
* Output routine called during the final phase;
#COUNT=#COUNT+1;
DO 10 FOR #2=#COUNT-1 ;
WRITE 1=10, ((E14,6)) TIME, CycBact, VOBS<0,#2> ;
WRITE 1=11, ((E14,6)) TIME, CycBact, VOBS<0,#2> ;
WRITE 1=12, ((E14,6)) TIME, Ln, Lc, Hc % ;
WRITE 1=13, ((E14,6)) TIME, Ln, Lc, Hc % ;
WRITE 1=14, ((E14,6)) TIME, H2AX % ;
WRITE 1=15, ((E14,6)) TIME, H2AX % ;
WRITE 1=16, ((E14,6)) TIME, Lm, Hm, Le, He % ;
WRITE 1=17, ((E14,6)) TIME, Lm, Hm, Le, He % ;
LABEL 10 ;
**;

SETNOFIT;
WHENEVER TIME = TOBS % CALL PRINT;
**;

BEGIN;
STOP;
```